

The Cognitive Wireless Era: AI/ML as the Engine of Next Generation Communication Networks

I.F. AKYILDIZ

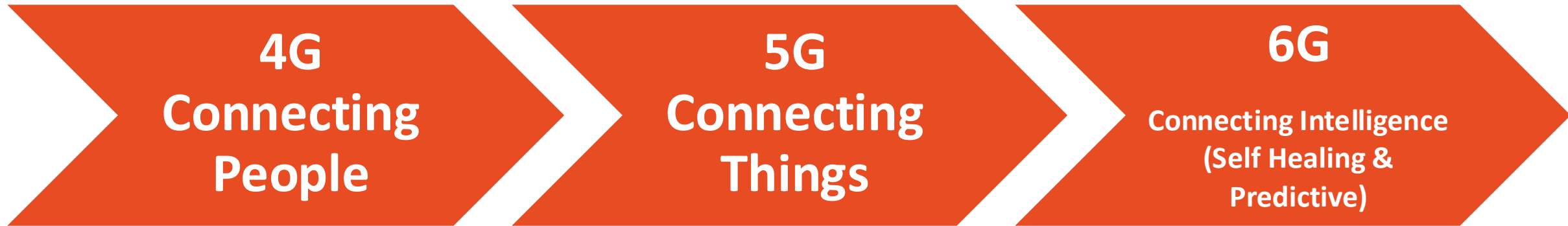
Georgia Institute of Technology

ian@ece.gatech.edu

International Telecommunication Union (ITU)

ianaky@itu.int

Brink of a New Era: From Reactive to Cognitive Networks



AI is not as a tool, but the *fundamental fabric* of this new, cognitive network.

"Cognitive Era" will transform society, healthcare, industry, and our relationship with technology.

WHY IS THIS TRANSITION CRUCIAL?



- **Economic Impact:** Trillions in global GDP

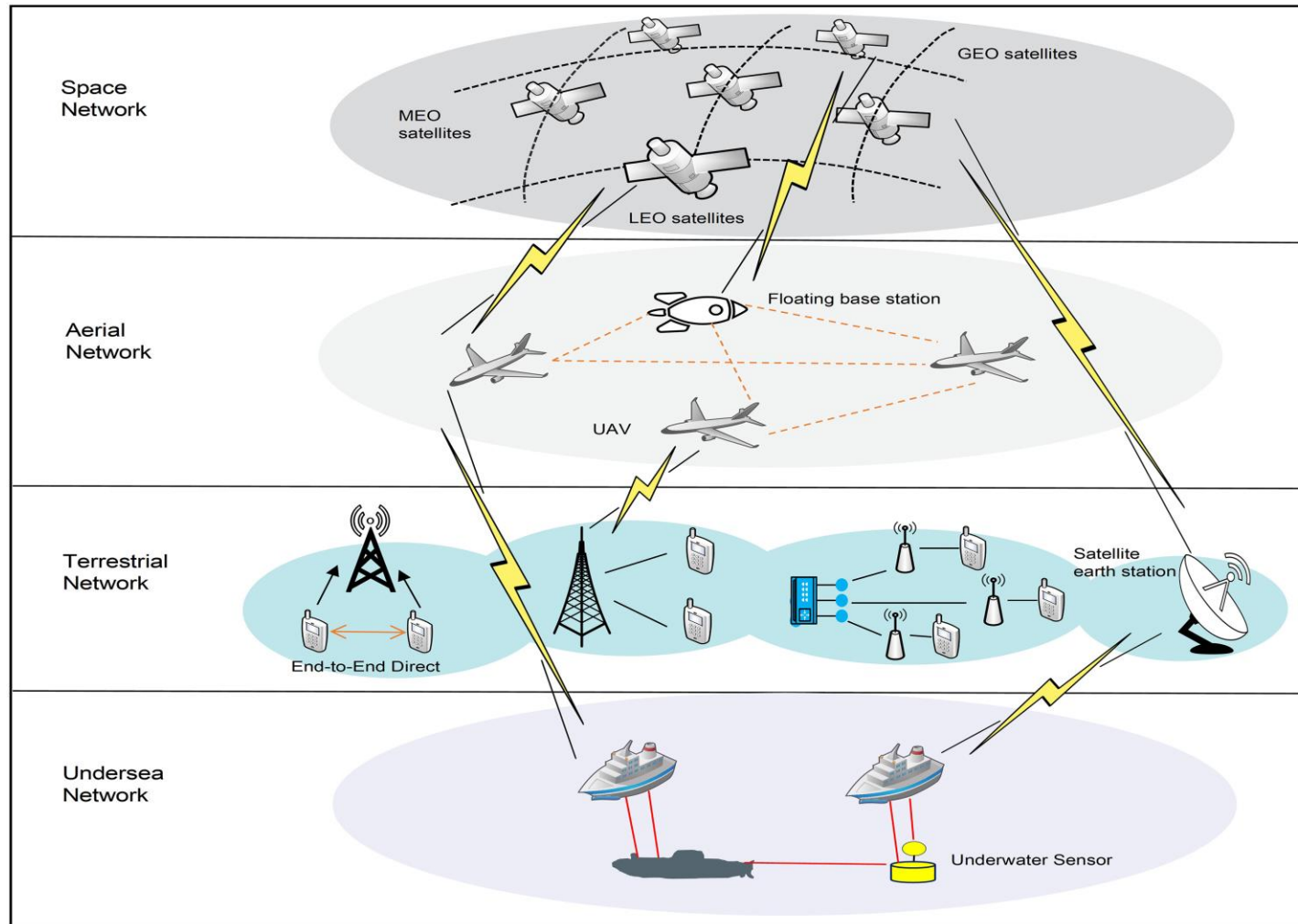


- **Societal Impact:** Remote surgery, autonomous transportation, disaster response, environmental monitoring



- **Geopolitical Impact:** Leadership in 6G is leadership in the future digital economy.

GLOBAL NETWORK 3D COVERAGE



NETWORKING 2030-2040

Applications



Telepresence

- Conference
- Tourism
- Remote Assistance



Healthcare

- Remote Surgery
- Emergency Resppnse
- Diagnosis & Patients Files



Education

- Remote Learning
- Virtual Labs



Automotive

- Autonomous Driving
- Driver Assistant
- Logistics
- Traffic Management



Entertainment

- Gaming
- Sports Broadcast



Industrial Automation

- Manufacturing
- Robots
- UAVs/DRONES

NETWORK

6G/Beyond; ALL PARADIGMS (GLOBAL COVERAGE; 3D)

FUNDAMENTAL DESIGN PRINCIPLE

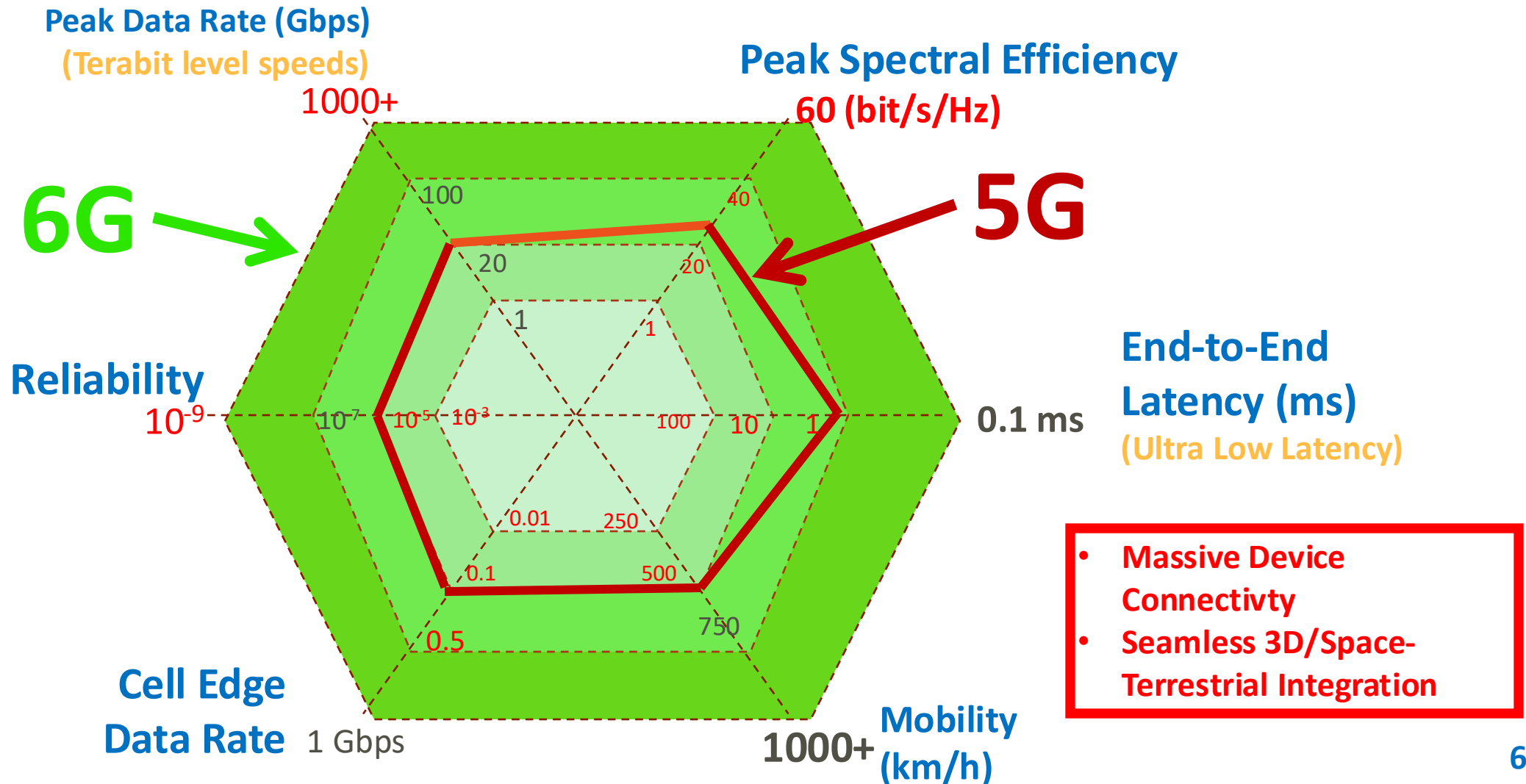
AI & ML

SECURITY/PRIVACY/SAFETY

EVOLUTION FROM 5G TO 6G

I. F. Akyildiz, A. Kak, S. Nie,

“6G AND BEYOND: THE FUTURE OF WIRELESS COMMUNICATIONS SYSTEMS”,
IEEE Access Journal, July 2020.

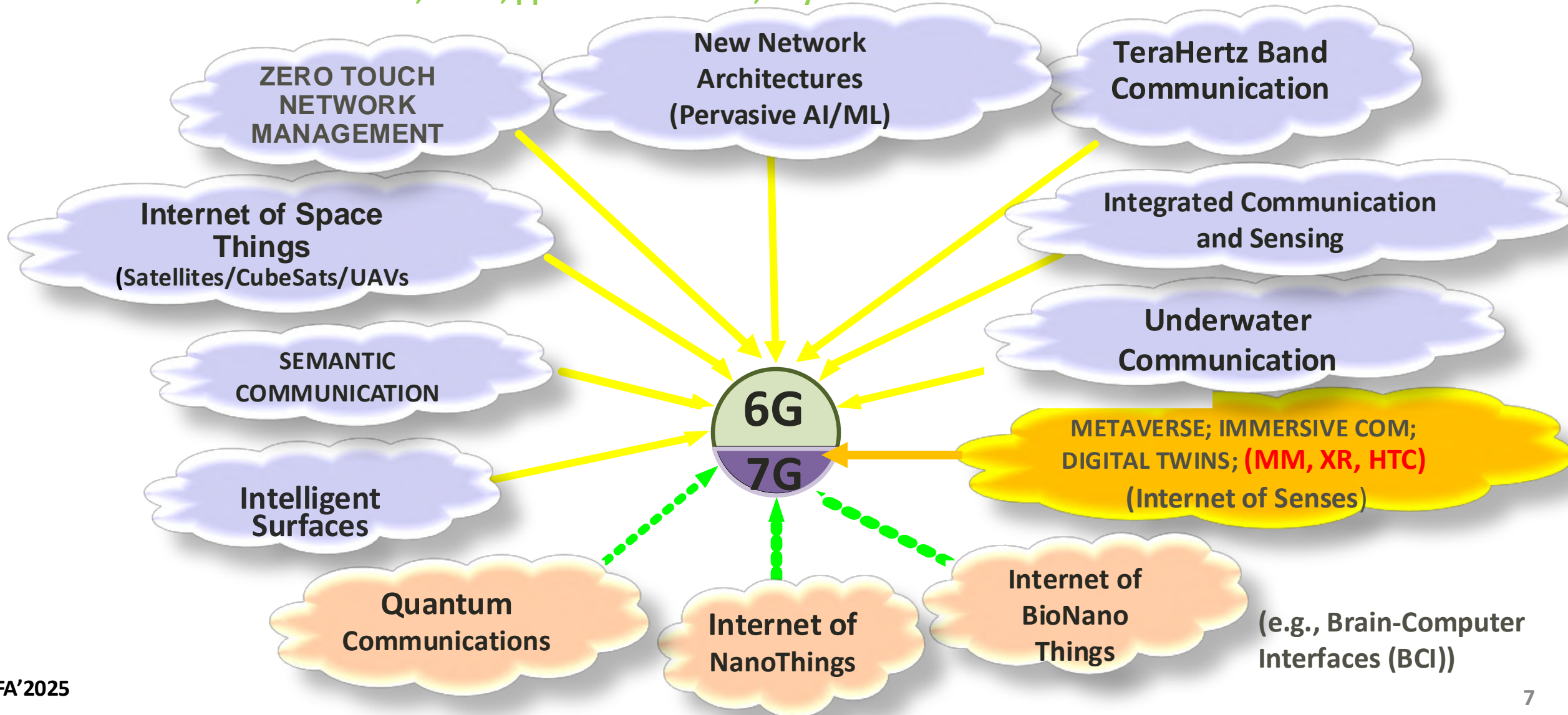


Key Enabling Technologies for 6G and BEYOND

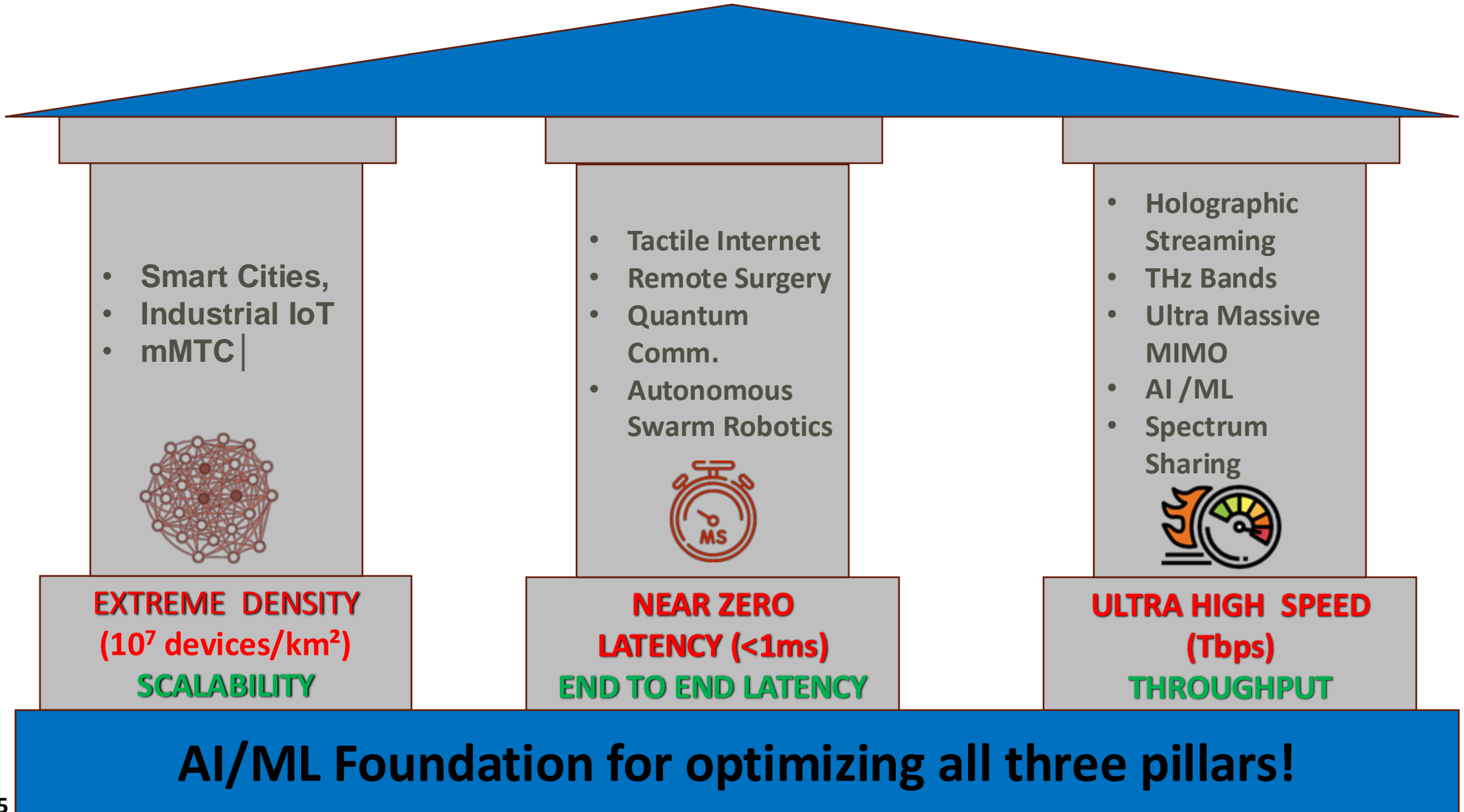
I. F. Akyildiz, A. Kak, S. Nie

“6G AND BEYOND: THE FUTURE OF Wireless Communication Systems”,

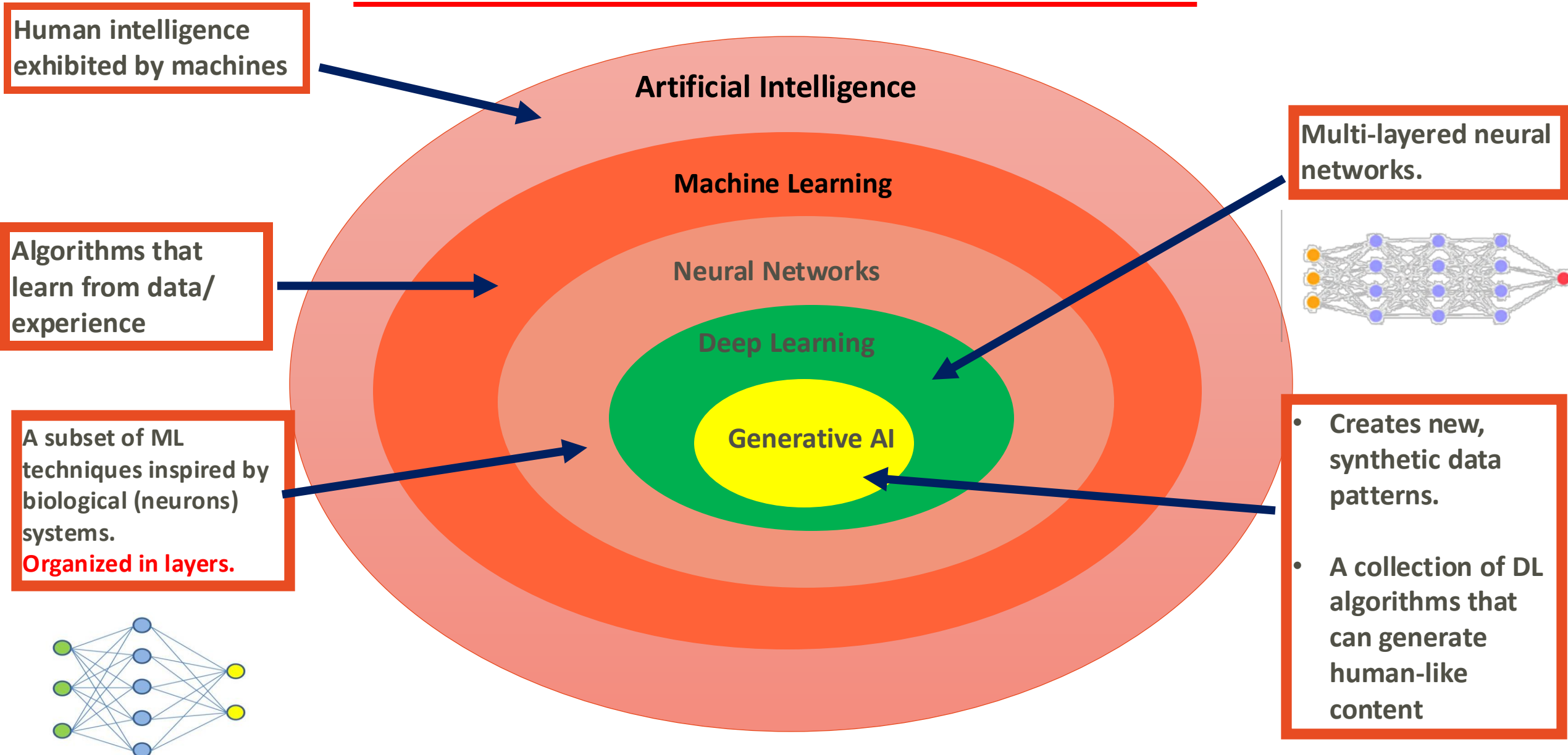
IEEE Access Journal, Vol. 8, pp. 133995-134039, July 2020.



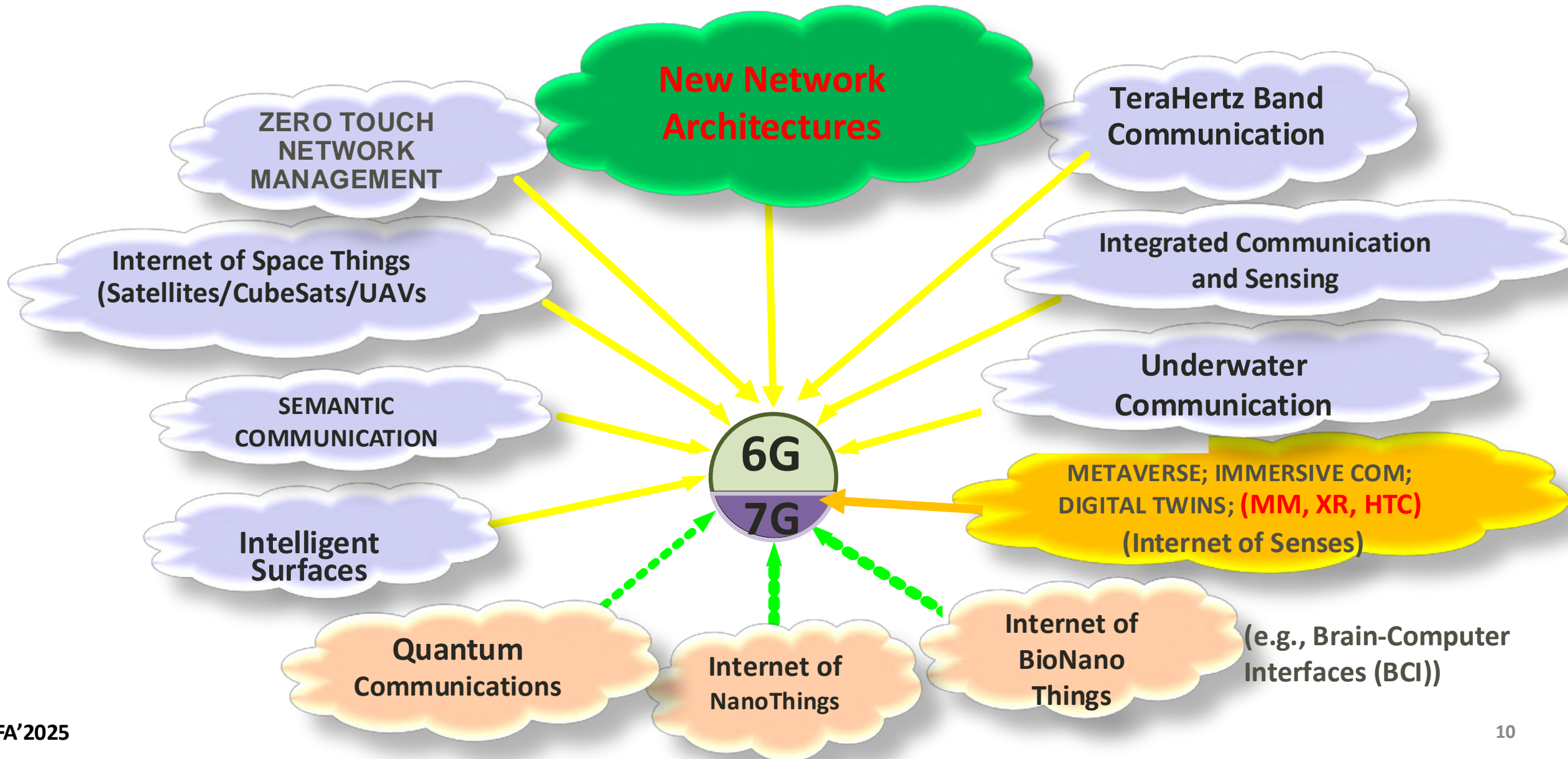
FUTURE WIRELESS SYSTEMS VISION



AI vs ML vs NN vs DL vs G-AI



AI-Native as the Engine of NG Wireless Systems



CASE 1: The AI-Native Imperative: A Paradigm Shift

Why Old-School Methods Fail

- Fixed rules and manual tuning
- Classical solutions struggle with **scale, nonlinearity, non-stationary, partial observability and real-time constraints**
- Future networks are **dynamic, large-scale, orders of magnitude more complex, very fast, and highly unpredictable**

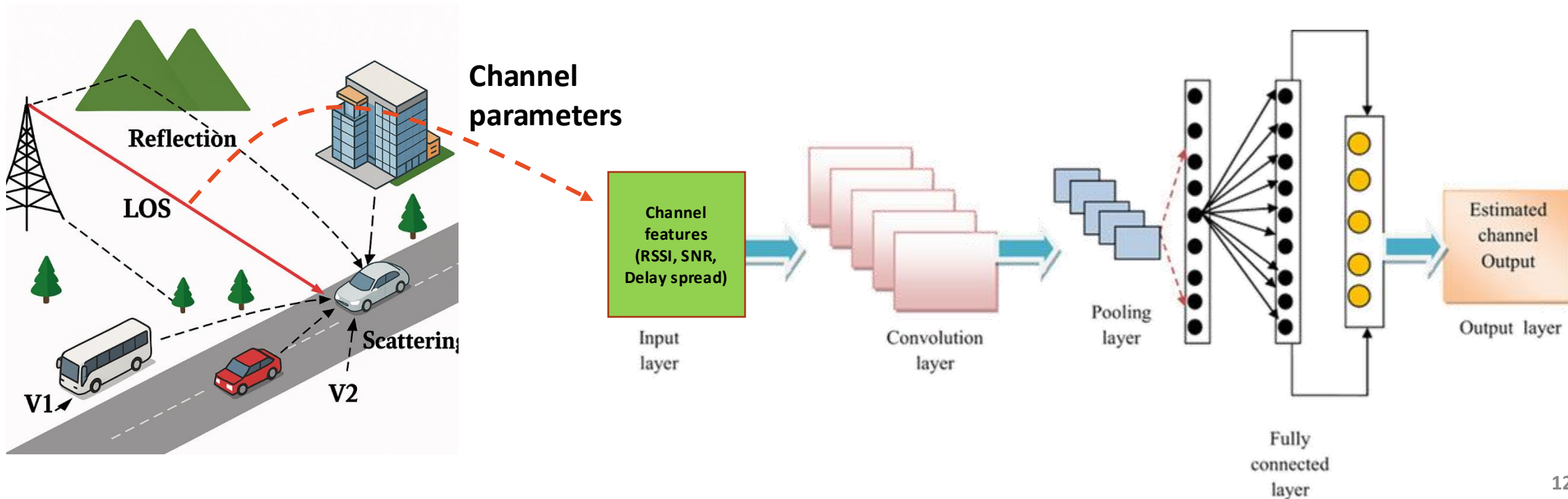
How AI-Native Can Help

- AI is not an external add-on but is embedded within the functional fabric of all network layers **(PHY, MAC, Network, Higher Layers)**.
- Enables real-time adaptation, handles massive data, and makes autonomous decisions.
- Delivers NWs that are smarter, more efficient, and more secure.
- AI-Native brings **adaptability, speed, and intelligence making future networks smarter, greener, and more secure.**

Case 1: AI-Native PHY Layer Research Challenges

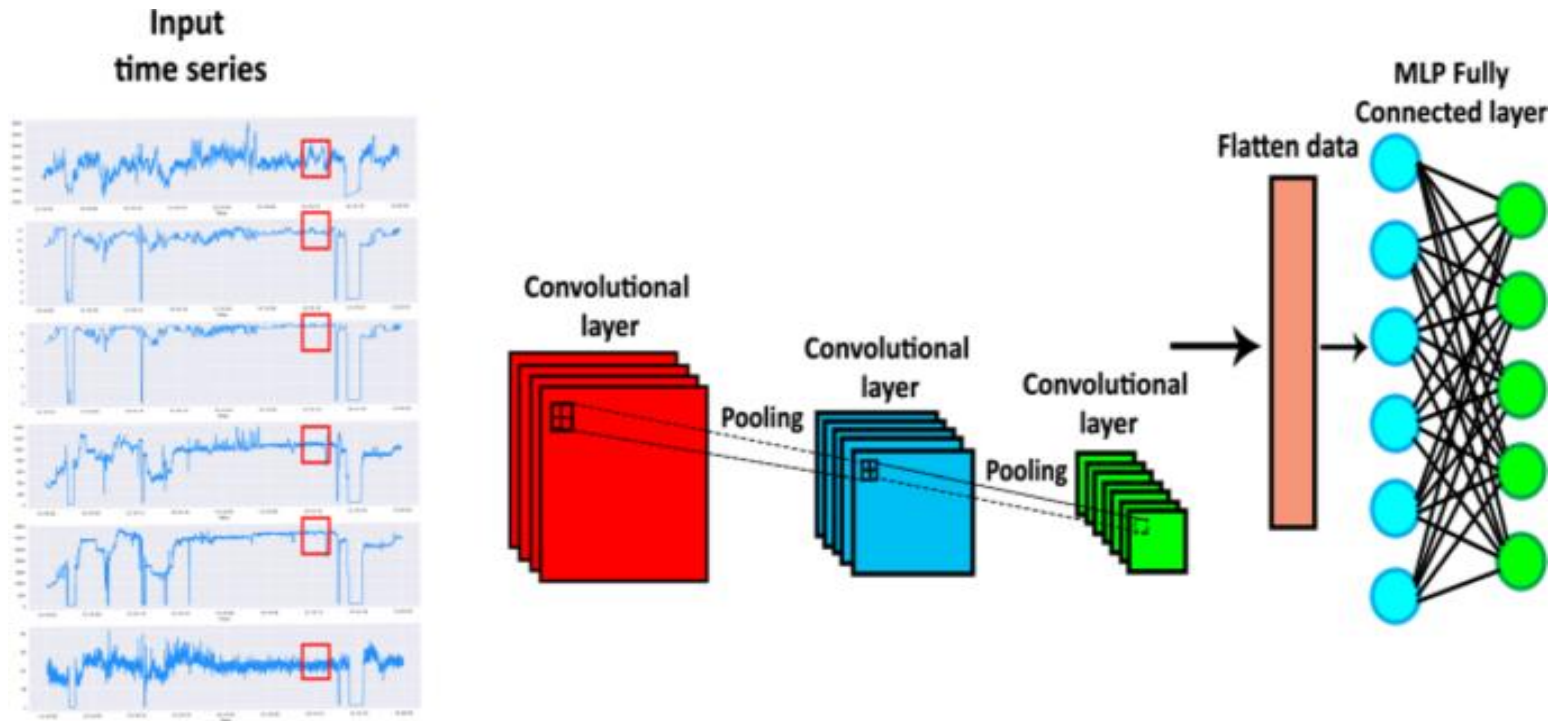
Channel Estimation & Optimal MCS Selection

- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs) (e.g., LSTMs)
- Transformer Models



Case 1: AI-Native PHY Layer Research Challenges

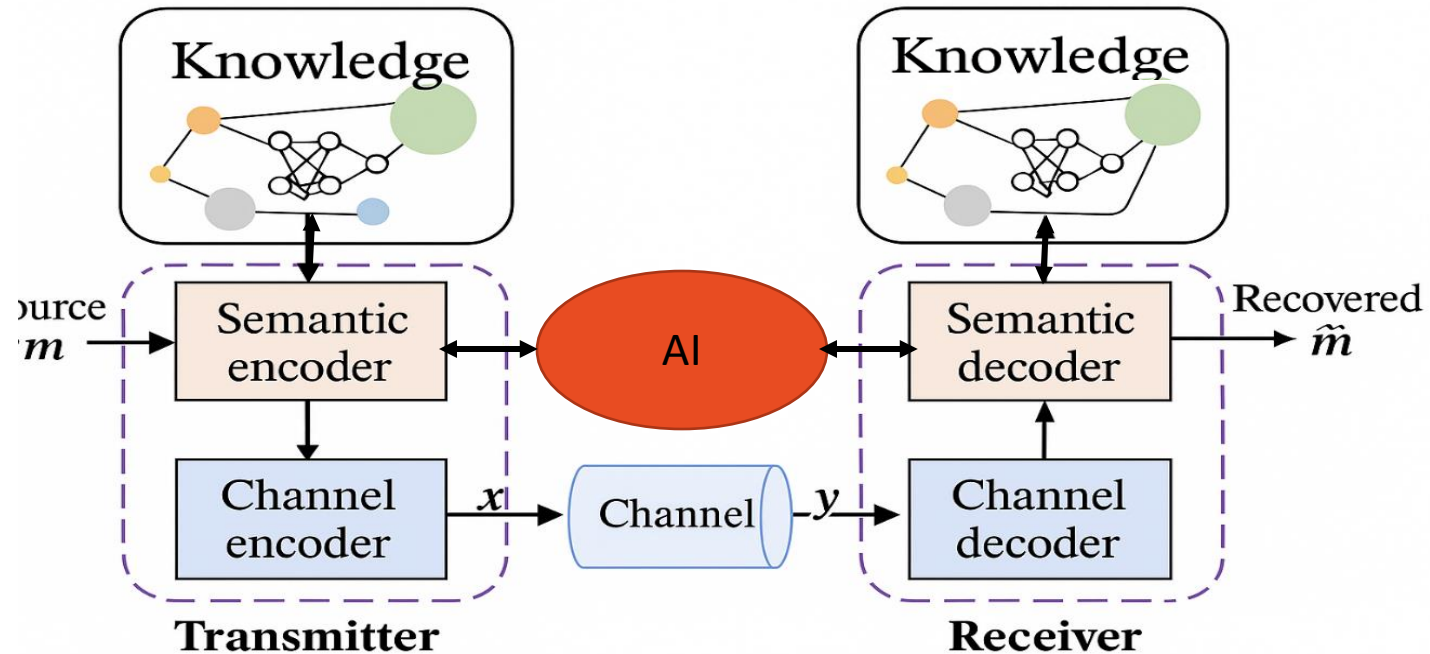
Time-series Signal Processing Hybrid CNN-RNN models



Case 1: AI-Native PHY Layer Research Challenges

Semantic Communications

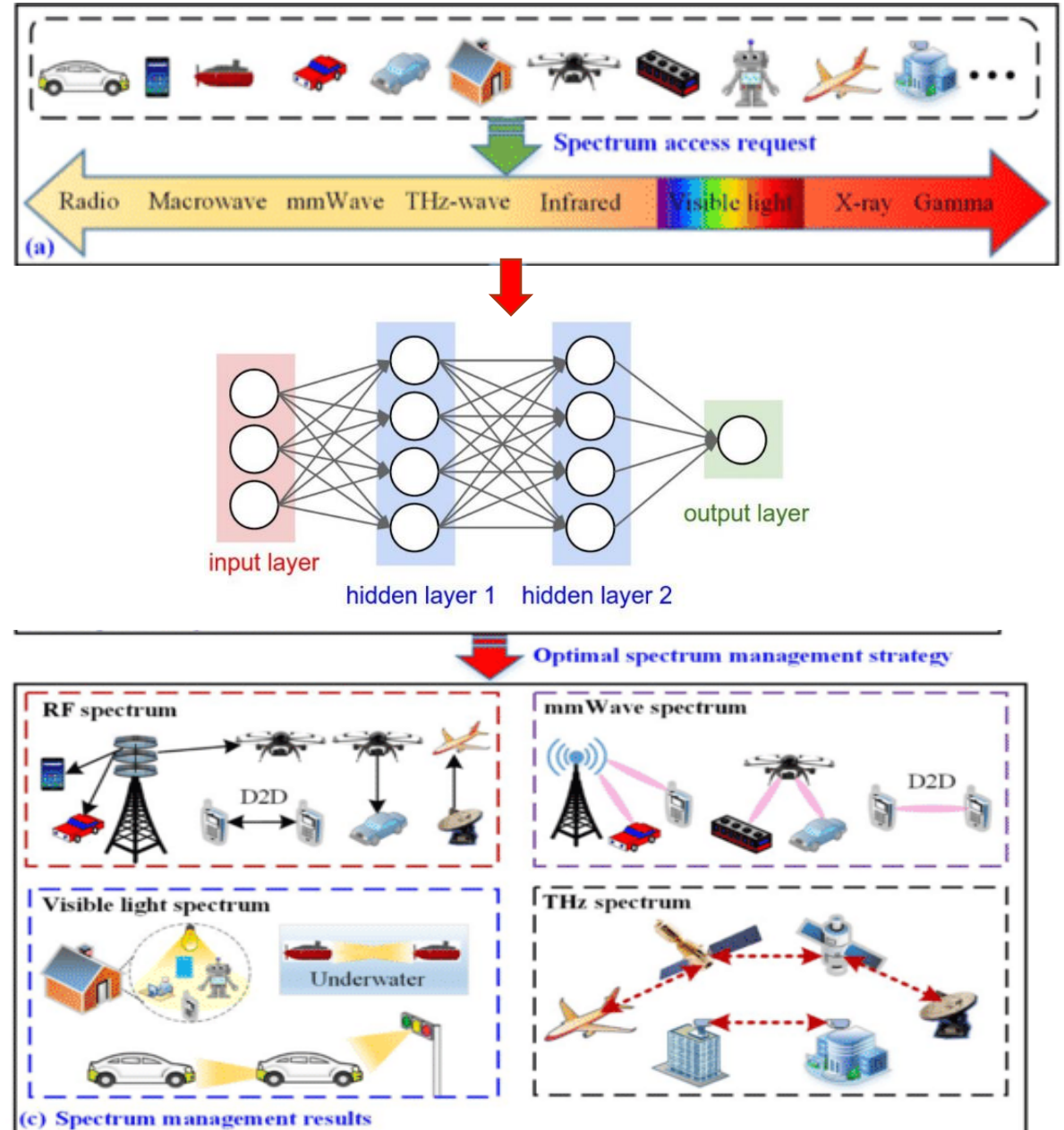
AI can extract meaning rather than just bits?



Case 1: AI-Native PHY Layer Research Challenges

Spectrum & Energy Efficiency

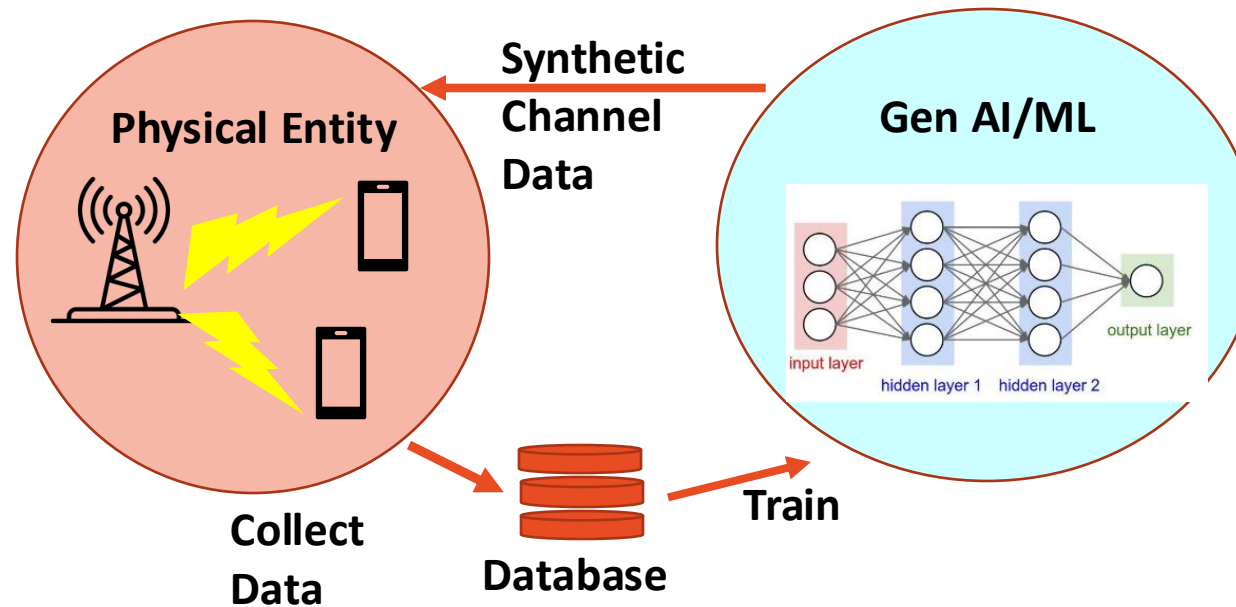
- DL-based Spectrum Sensing & Sharing
- Power-Aware ML (predictive energy management for IoT devices)



Case 1: AI-Native PHY Layer Research Challenges

AI-Generated Wireless Signals

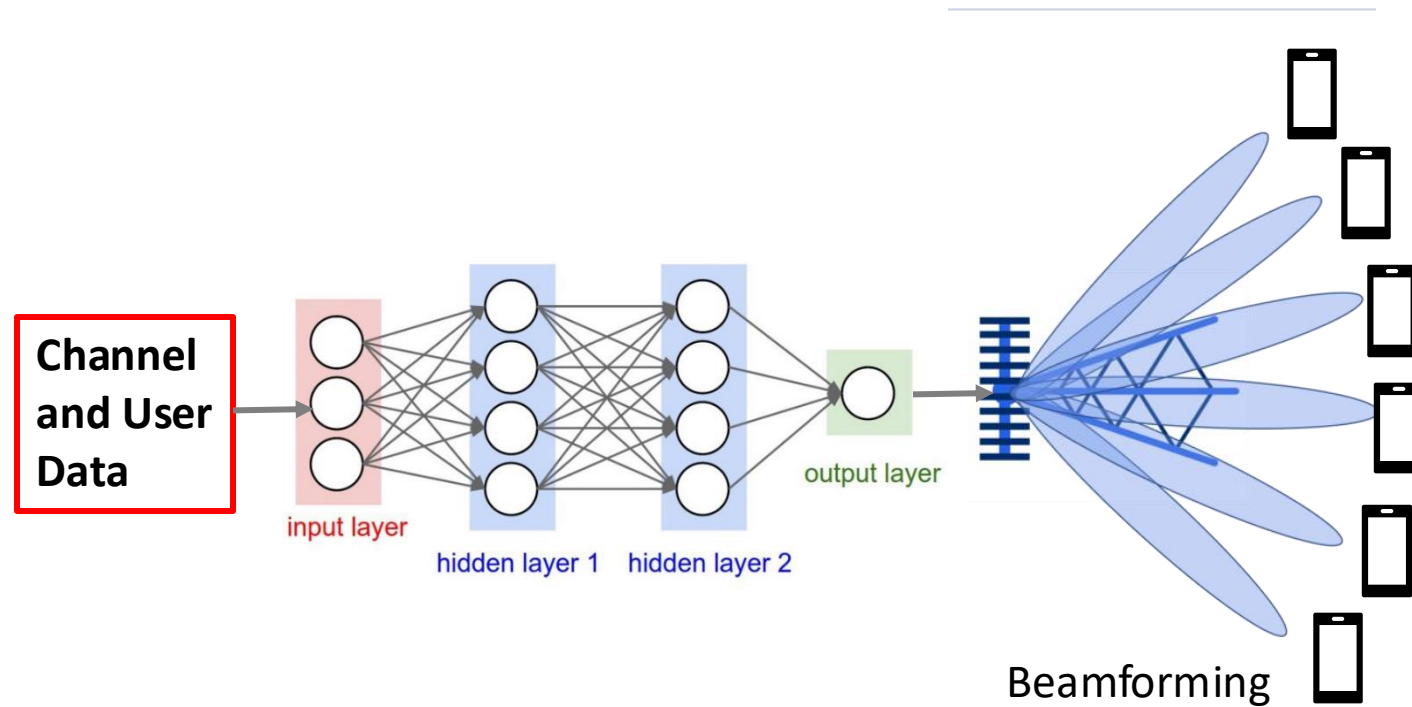
GANs (Generative Adversarial Networks) may create synthetic but realistic channel data?



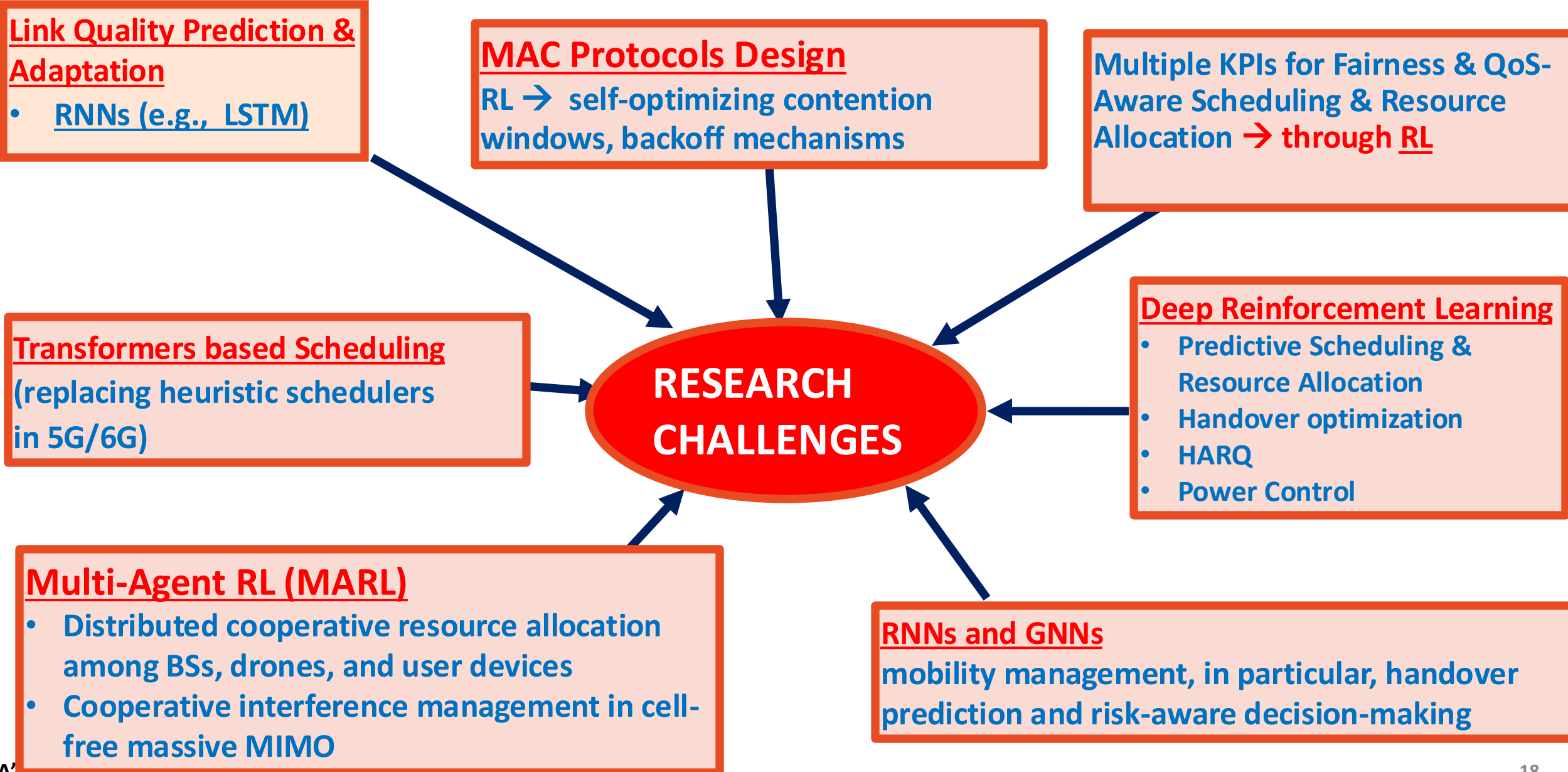
Case 1: AI-Native PHY Layer Research Challenges

MIMO & Beamforming & Beam Management

- RNNs for MIMO
- RL for dynamic beam alignment
- FL for distributed beam optimization



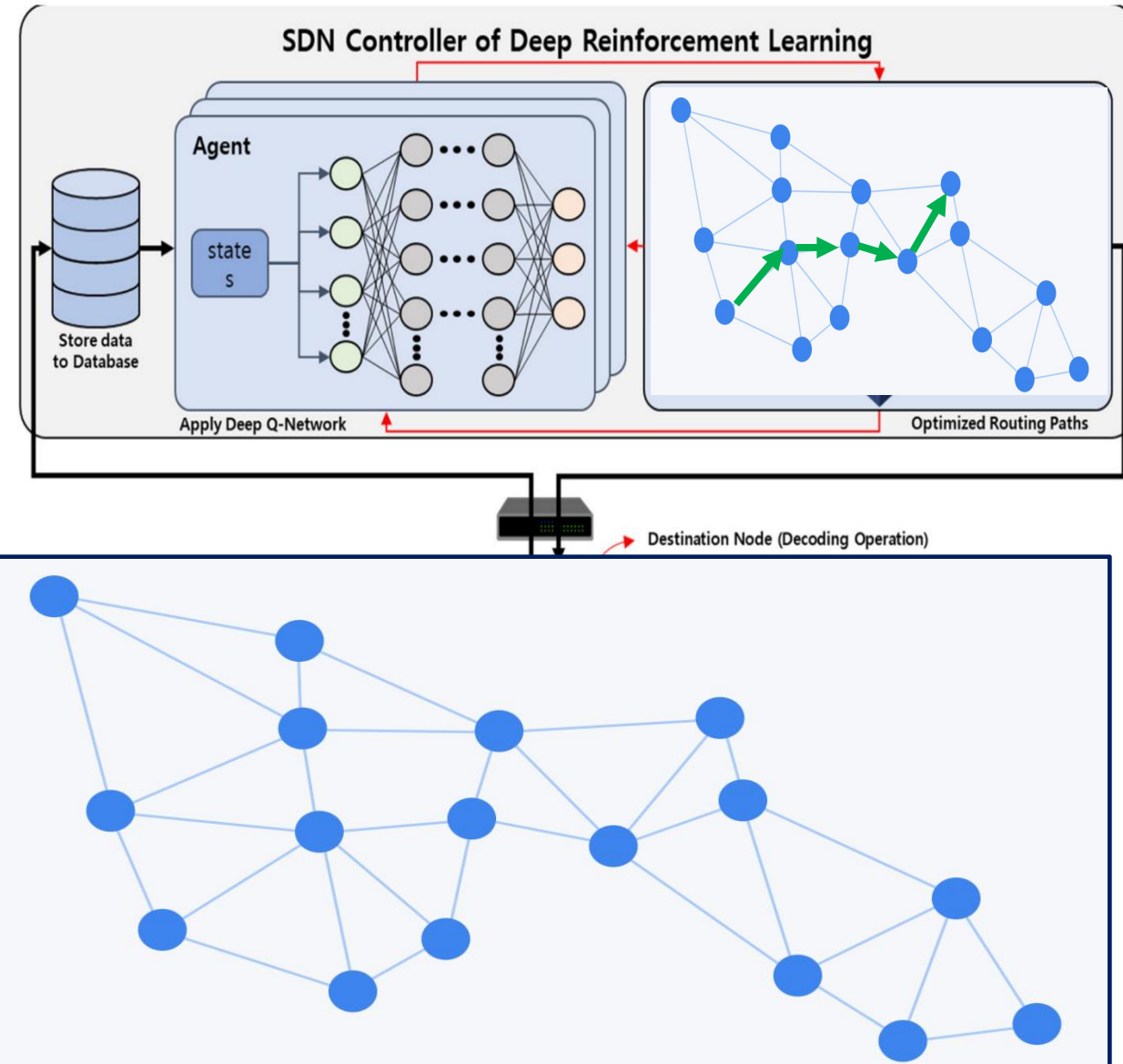
AI-Native MAC Layer



AI-Native NW (ROUTING) Layer Research Challenges

DRL for Adaptive Routing:

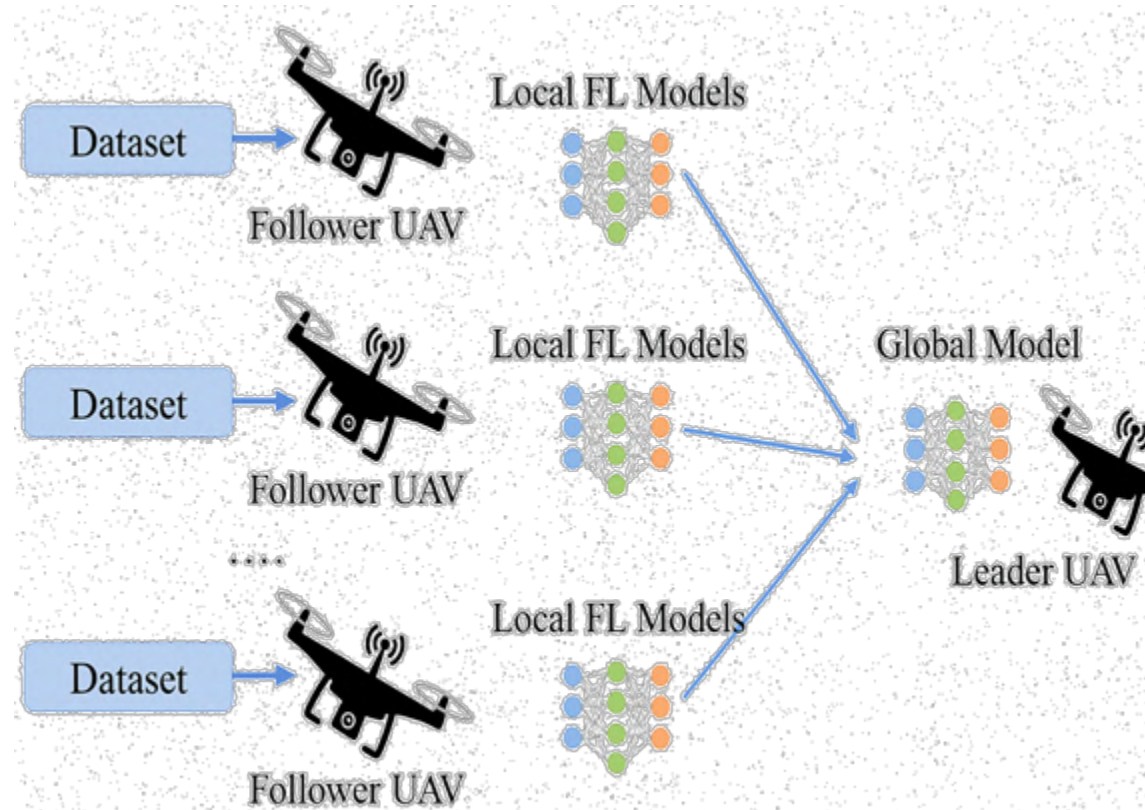
DRL agents at nodes learn optimal paths based on real-time conditions (latency, congestion, energy)



AI-Native NW (ROUTING) Layer Research Challenges

Federated Learning (FL):

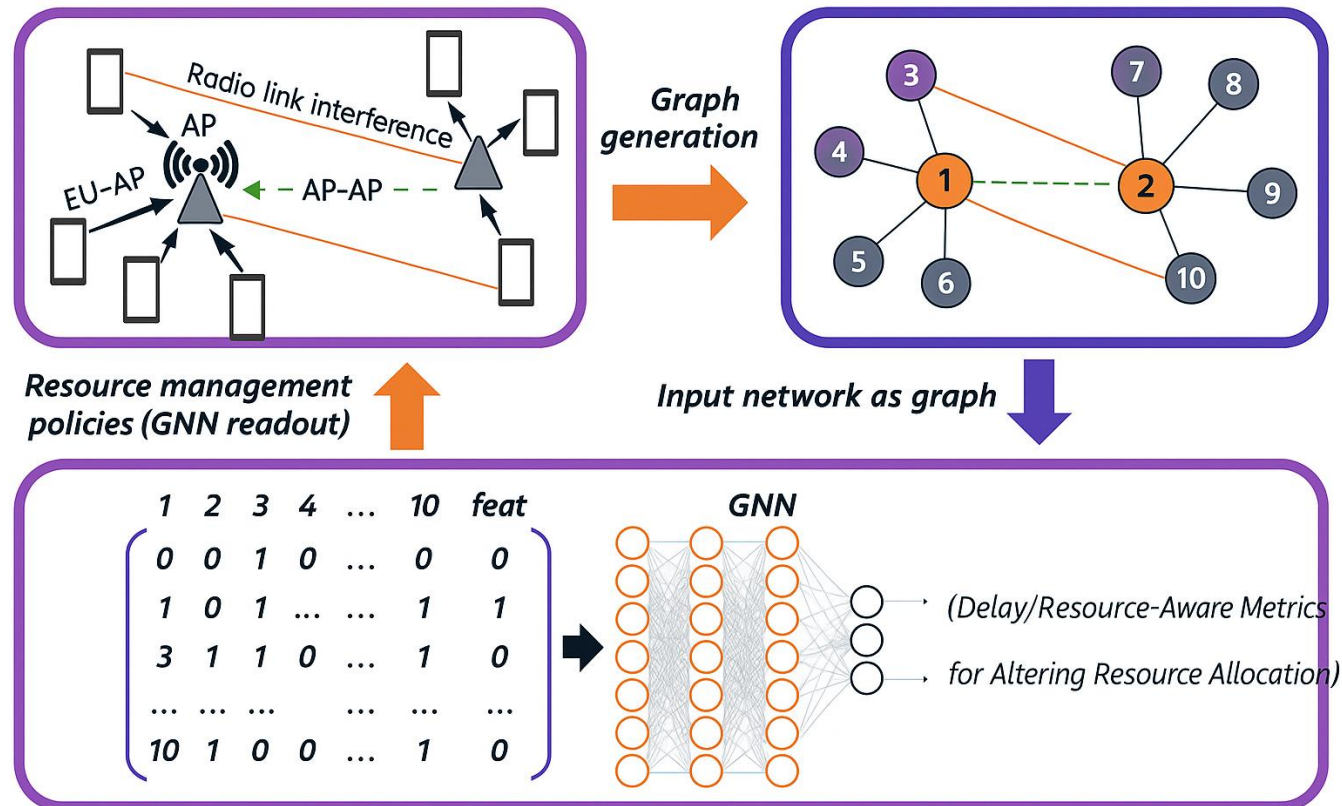
Enables collaborative training of routing models across distributed nodes (drones, vehicles) without sharing raw data, preserving privacy and scalability.



AI-Native NW (ROUTING) Layer Research Challenges

Graph Neural Networks (GNNs):

- Model network topology.
- Used for global path computation, congestion prediction, and dynamic network slicing by understanding relationships between nodes.



THE SCALABILITY PROBLEM

Can we build AI that is fast and efficient enough for the real world?

Real Time & Resource Constraints:

Running complex AI models (GPT) on resource-constrained devices (sensors), e.g., URRLC (<1ms)

Solutions:

- Edge computing with lightweight models (TinyML)
- Neuromorphic Computing

Data Scarcity & Generalization:

AI models require very large datasets but struggle to generalize

Solution:

Techniques like transfer learning, meta-learning, and synthetic data generation

Energy Efficiency:

- High energy costs of large models (e.g., Transformers).
- Need Green AI !!

Solution:

Develop sparsity-aware NNs for energy-efficient inference.

Hardware-Aware AI:

Co-design of algorithms & specialized HW for efficiency

RESEARCH
CHALLENGES

THE TRUST PROBLEM

Can we trust the decisions these AI systems make?

Explainable AI (XAI)

Develop Explainable AI (XAI) techniques to debug and justify the decisions made by opaque "black-box" AI models, employing a hybrid symbolic-neural approach for greater transparency.

Trustworthy AI for Security & Robustness:

- Protecting ML systems against adversarial attacks
- Support interpretability or beamforming and resource allocation.

Ethics & Fairness:

Addressing bias, fairness, and accountability in AI-driven systems

**RESEARCH
CHALLENGES**

Privacy:

Implement FL and differential privacy.

THE INTEGRATION PROBLEM

Can we weave this new intelligence into our existing world?

Lack of Information Theory

that integrates computation & learning and determines fundamental limits of AI-native communication

HYBRID AI APPROACHES:

Combining model-based optimization with data-driven learning

HYBRID ENVIRONMENTS

Developing AI solutions that work across urban, rural, UW, and space environments

Solution:

Use meta-learning for scalability & fast adaptation to new scenarios

Foundation Models:

Large Language Models pre-trained on massive, multi-modal network data

**RESEARCH
CHALLENGES**

Legacy System Coexistence:

Enabling 5G/6G to work with legacy 4G systems

Standardization & Regulation & Interoperability

A lack of standardized AI interfaces hinders model sharing across vendors.

CASE 2: THz BAND COMMUNICATION

I.F. Akyildiz, J. M. Jornet and C. Han,

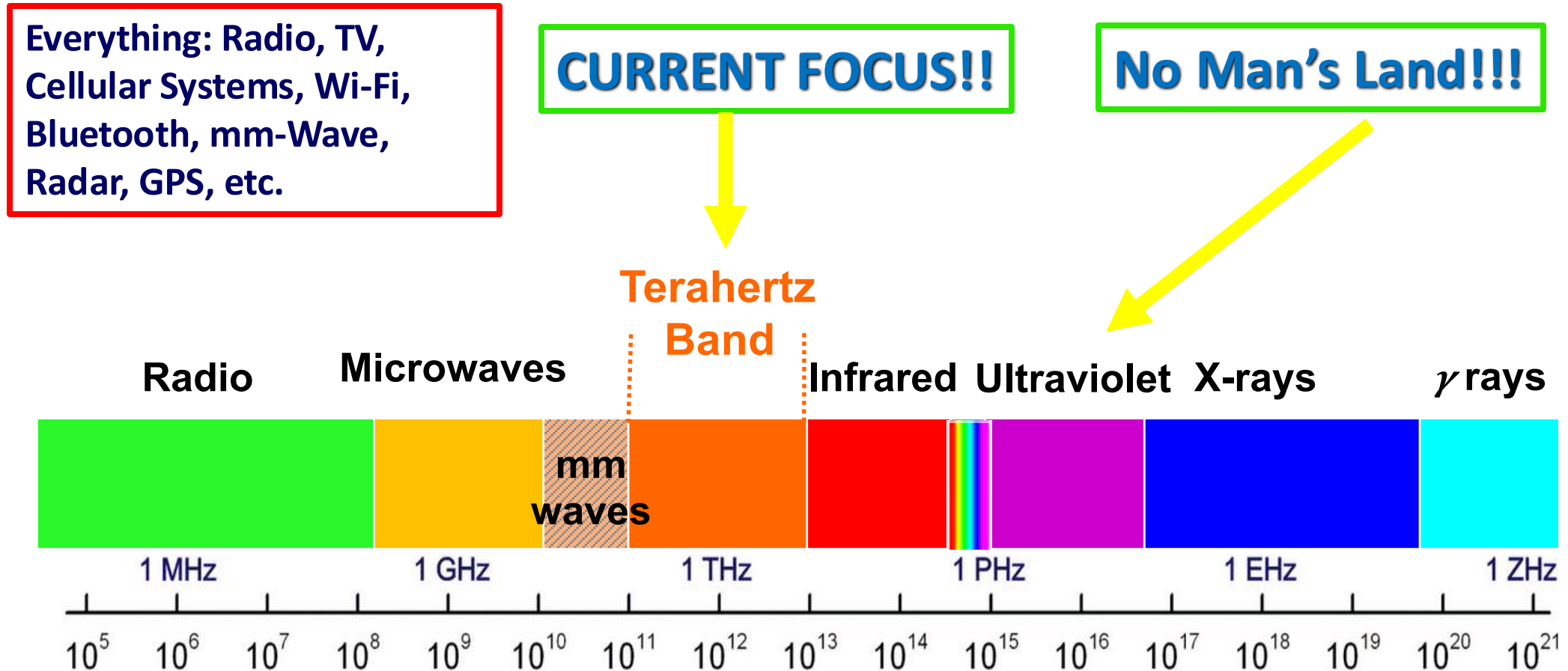
"TERANETS: ULTRA-BROADBAND COMMUNICATION NETWORKS IN THE TERAHERTZ BAND,"
IEEE WIRELESS COMMUNICATIONS MAGAZINE, VOL. 21, NO. 4, PP. 130-135, AUGUST 2014.

I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet,

"TeraHertz Band Communication: An Old Problem Revisited and Research Directions for the Next Decade", **IEEE Transactions on Communications, June 2022.**

- **6G REQUIREMENTS** (Min End to End Latency; Very High Reliability; Very High Data Rates)
- **Exponential growth of wireless data traffic:**
 - **More Devices → Multi-billion fixed-mobile-connected devices by 2025**
 - **Faster Connections → Wireless data rates have doubled every 18 months over the last three decades**
 - **Wireless Terabit-per-second (Tbps) links will become a reality within the next 5 years → HOW??? → Explore high frequencies !!**

CASE 2: THz BAND COMMUNICATION



TERANETS (formerly GRANET; 2008-2013):

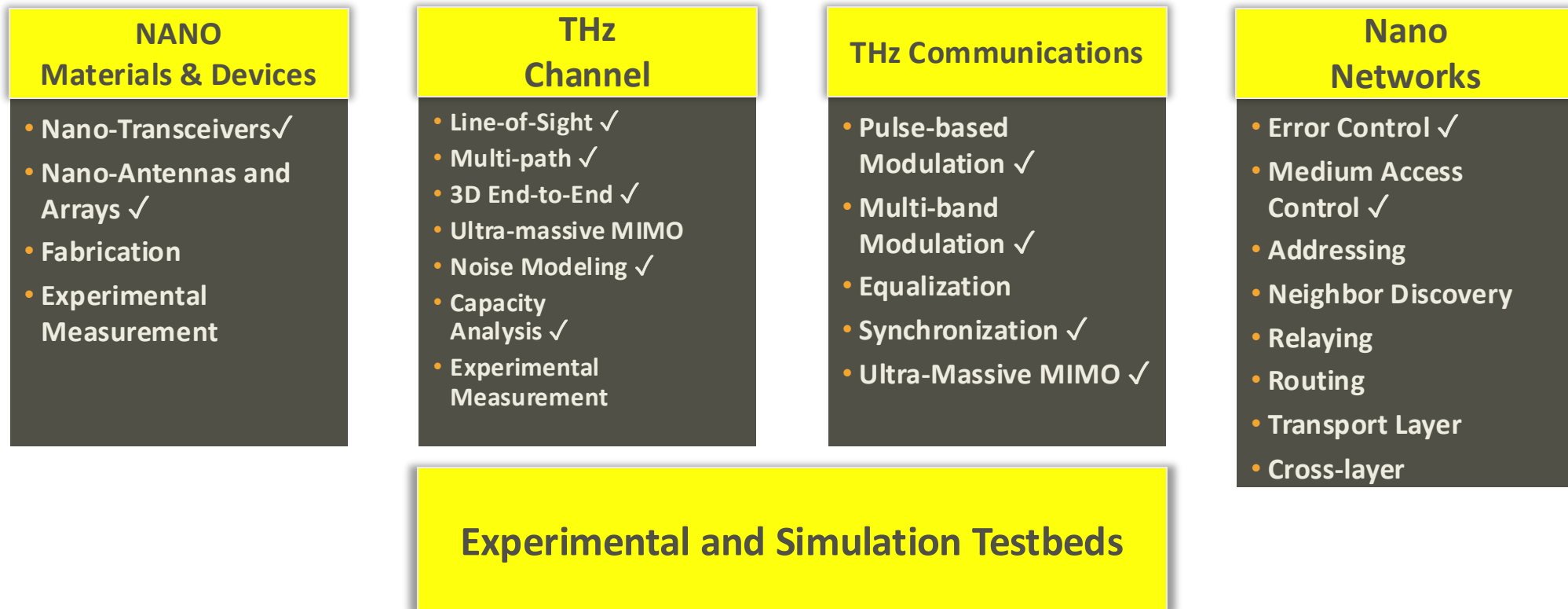
“GRAPHENE BASED NANO SCALE COMMUNICATION NETWORKS IN THZ BAND”

NSF; US ARMY; FiDiPro; CATALUNA; HUMBOLDT; KACST, etc..

2008-2013; 2013-2016 & 2016-2020 ; 2018-2022

• Objectives:

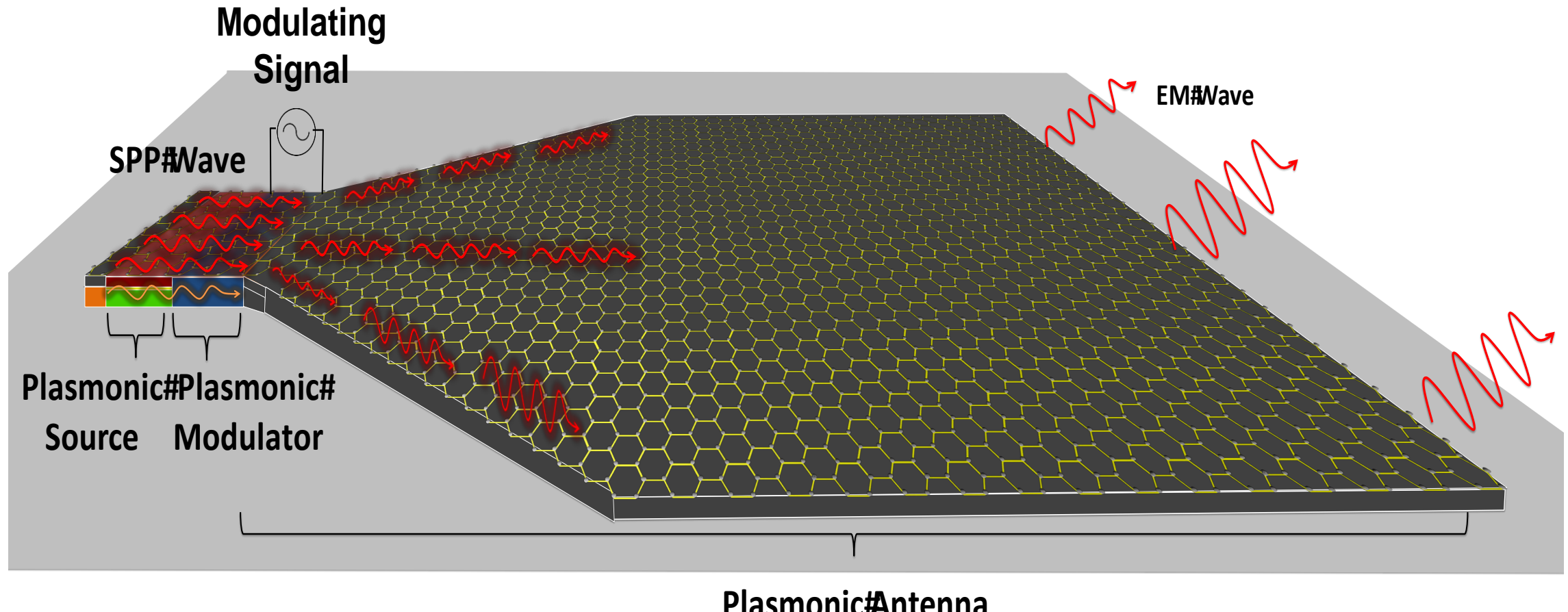
- **To demonstrate the feasibility of graphene-enabled EM communication**
- **To establish the theoretical foundations for EM nanonetwork**
- **To establish the theoretical and experimental foundations of ultra-broadband com nets in the (0.1-10) THz band**



TeraHertz Band Plasmonic Front-end

(TRANSCIVER+ANTENNA)

- I. F. Akyildiz & J. M. Jornet, "Graphene Based Plasmonic Nano-Antenna for THz Band Com in NanoNetworks"
U.S. Patent No. 9,643,841, May 9, 2017.
- I. F. Akyildiz & J. M. Jornet, "Graphene Plasmonic Nano-transceiver for Wireless Com in the THz Band,"
U.S. Patent No. 9,397,758, July 19, 2016.



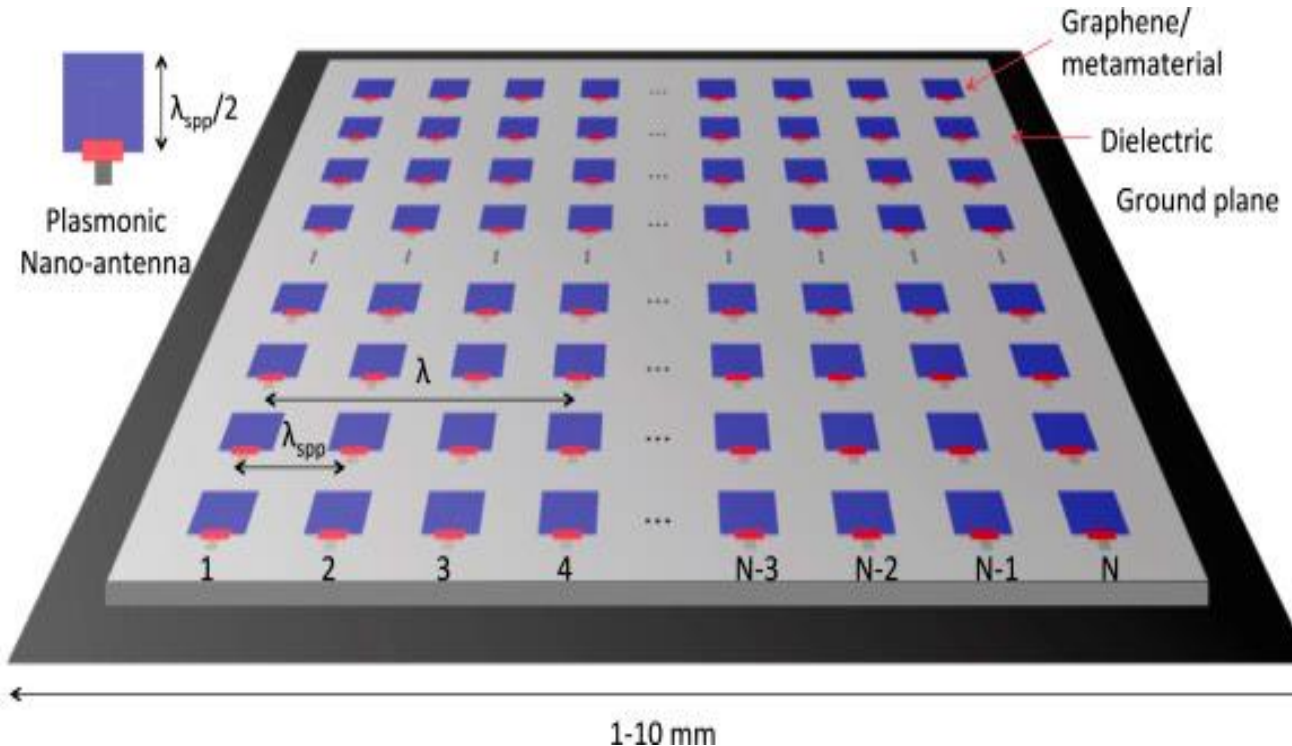
COMBATING DISTANCE PROBLEM: ULTRA-MASSIVE MIMO

I. F. Akyildiz and J. M. Jornet

“Realizing Ultra-Massive MIMO Communication in the (0.06–10) TeraHertz Band”

U.S. Patent 15/211,503 awarded on Sept. 7, 2017.

Planar Array with 32x32 antenna elements in total of 1024 elements



UM-MIMO can achieve at least 10-fold increase in transmission distance at 300 GHz and 1 THz compared to M-MIMO

HYPERSURFACES: Programmable (SOFTWARE DEFINED) METASURFACES

A. Pitsillides, C. Liaskos, A. Tsioliaridou, S. Ioannidis, I. F. Akyildiz,

“Wireless Communication Paradigm Realizing Programmable Wireless Environments through Software-controlled Metasurfaces”

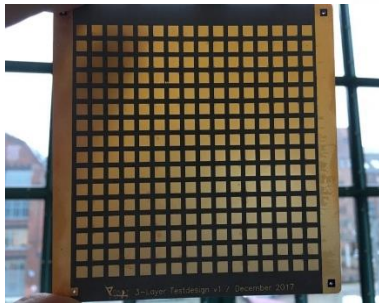
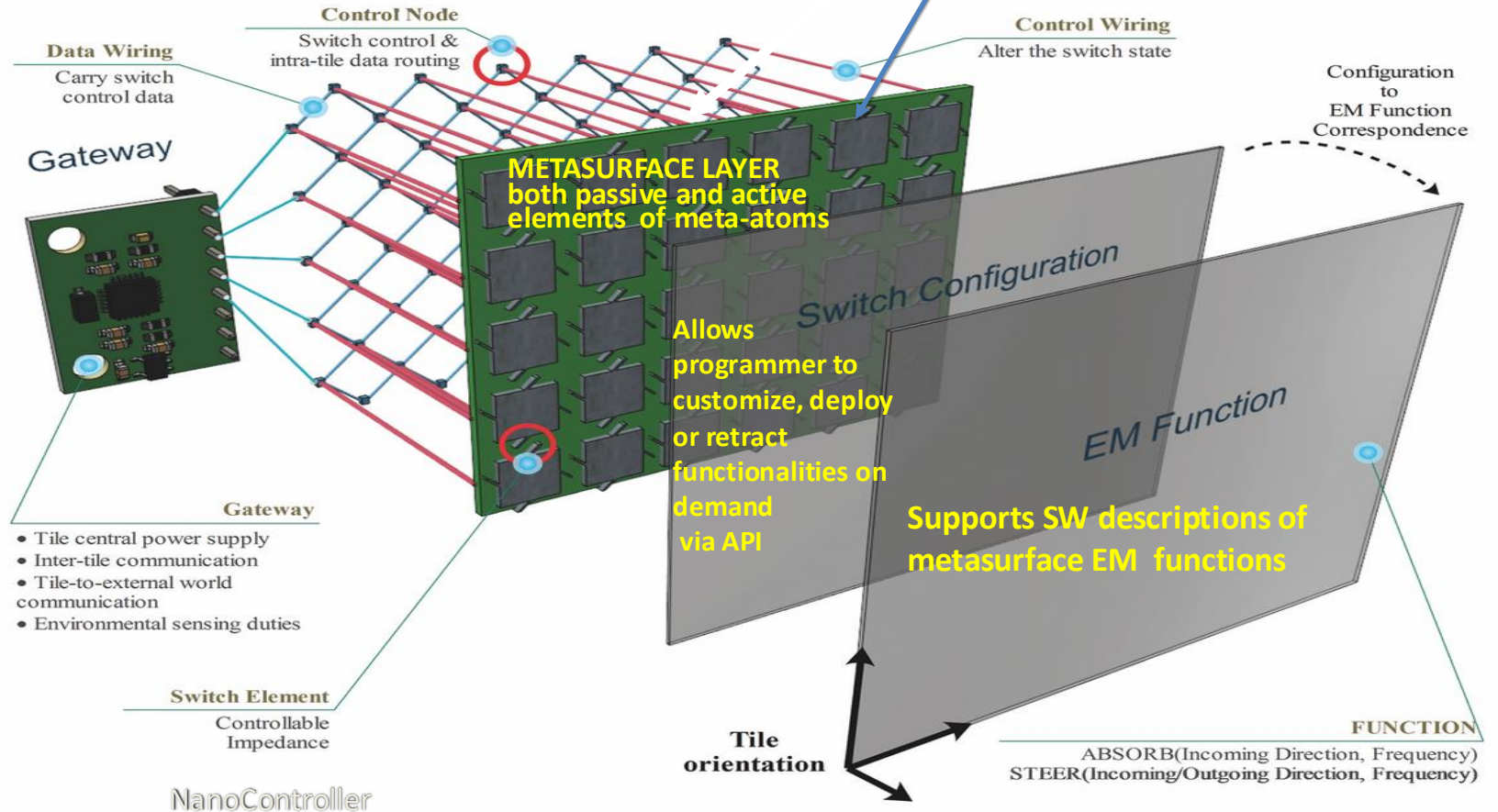
US PATENT, 10.547.116 B2; January 28, 2020.

meta atoms/
metallic
patches/
unit cells

EU-FET VISORSURF (2016-2021)

<http://www.visorsurf.eu>

8M Euro



CASE 2: AI-Native THz BAND COMMUNICATION

1. Channel Estimation & Beam Alignment

Problem: THz channels suffer from high path loss, molecular absorption, and dynamic blockage, requiring ultra-fast and accurate beamforming.

Best ML Solutions

```
graph TD; A[Best ML Solutions] --> B[DL (CNN, Transformer Models)]; A --> C[RL]; A --> D[FL]; B --> B1["CNN → Extract spatial features from THz channel measurements"]; B --> B2["Vision Transformers → Handle sparse THz channel matrices efficiently"]; C --> C1["Deep Deterministic Policy Gradient (DDPG) → Optimize beam alignment in real-time"]; C --> C2["Hierarchical RL → Multi-stage beam search for fast initial access"]; D --> D1["Enables distributed THz channel learning across multiple BSs without sharing raw data"];
```

DL (CNN, Transformer Models)

CNN → Extract spatial features from THz channel measurements

Vision Transformers → Handle sparse THz channel matrices efficiently

RL

Deep Deterministic Policy Gradient (DDPG) → Optimize beam alignment in real-time

Hierarchical RL → Multi-stage beam search for fast initial access

FL

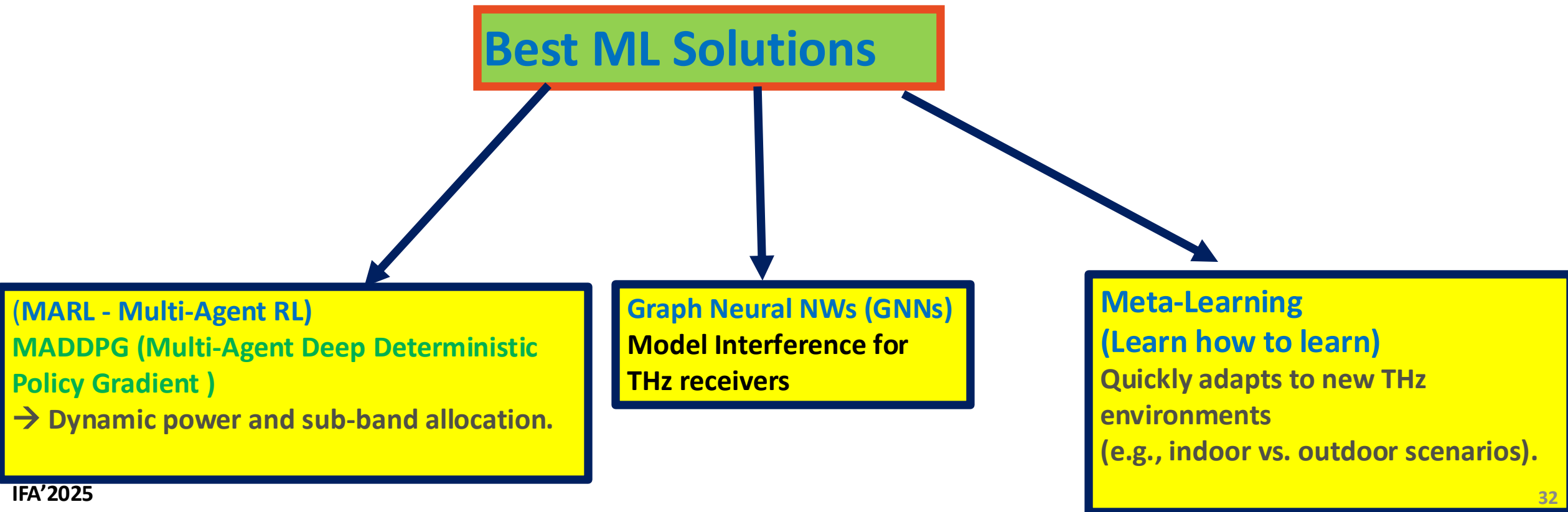
Enables distributed THz channel learning across multiple BSs without sharing raw data

CASE 2: AI-Native THz BAND COMMUNICATION

2. Resource Allocation & Spectrum Management

Problem:

THz bands have ultra-wide BW but suffer from dynamic interference and absorption peaks.



CASE 2: AI-Native THz BAND COMMUNICATION

3. Overcoming Molecular Absorption & Blockage

Problem: THz signals are absorbed by water vapor and oxygen molecules, leading to distance-dependent losses.

Best ML Solutions

```
graph TD; A[Best ML Solutions] --> B[Physics-Informed Neural Networks (PINNs)]; A --> C[Generative AI (Diffusion Models, GANs)]; A --> D[Bayesian Deep Learning];
```

Physics-Informed Neural Networks (PINNs)

Combines wave propagation physics with ML for better path loss prediction.

Generative AI (Diffusion Models, GANs)

Simulate THz channel variations under different humidity conditions.

Bayesian Deep Learning

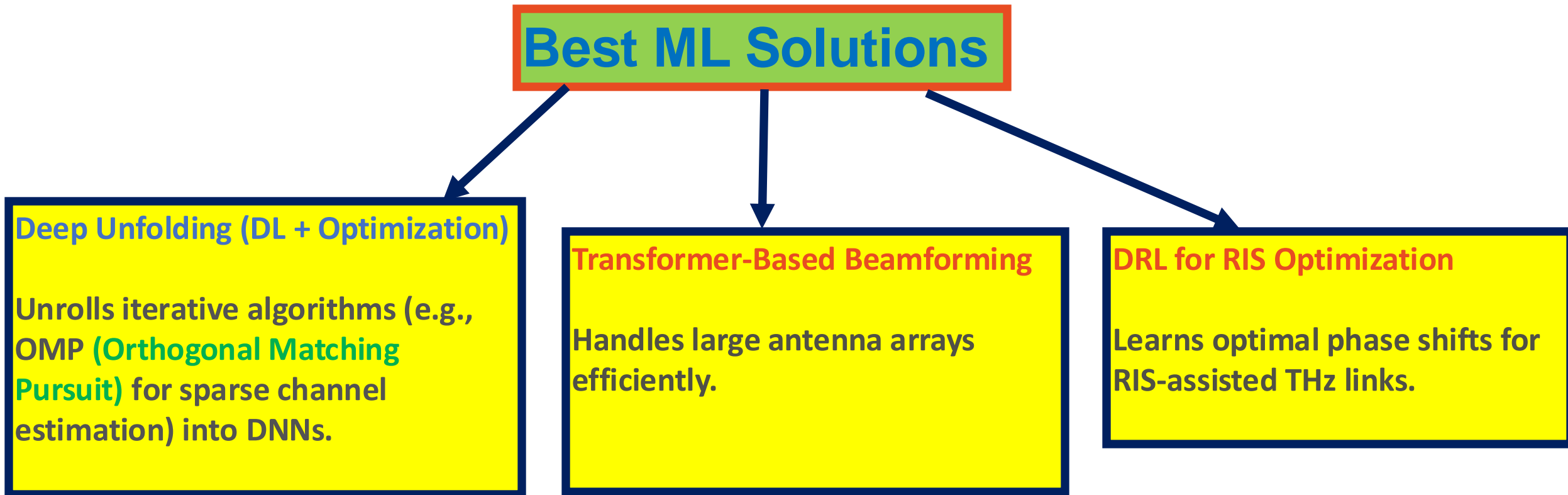
Quantifies uncertainty in THz link reliability.

CASE 2: AI-Native THz BAND COMMUNICATION

4. Ultra-Massive MIMO & Reconfigurable Intelligent Surfaces (RIS)

Problem:

THz requires ultra-massive antenna arrays, but traditional signal processing is too complex.

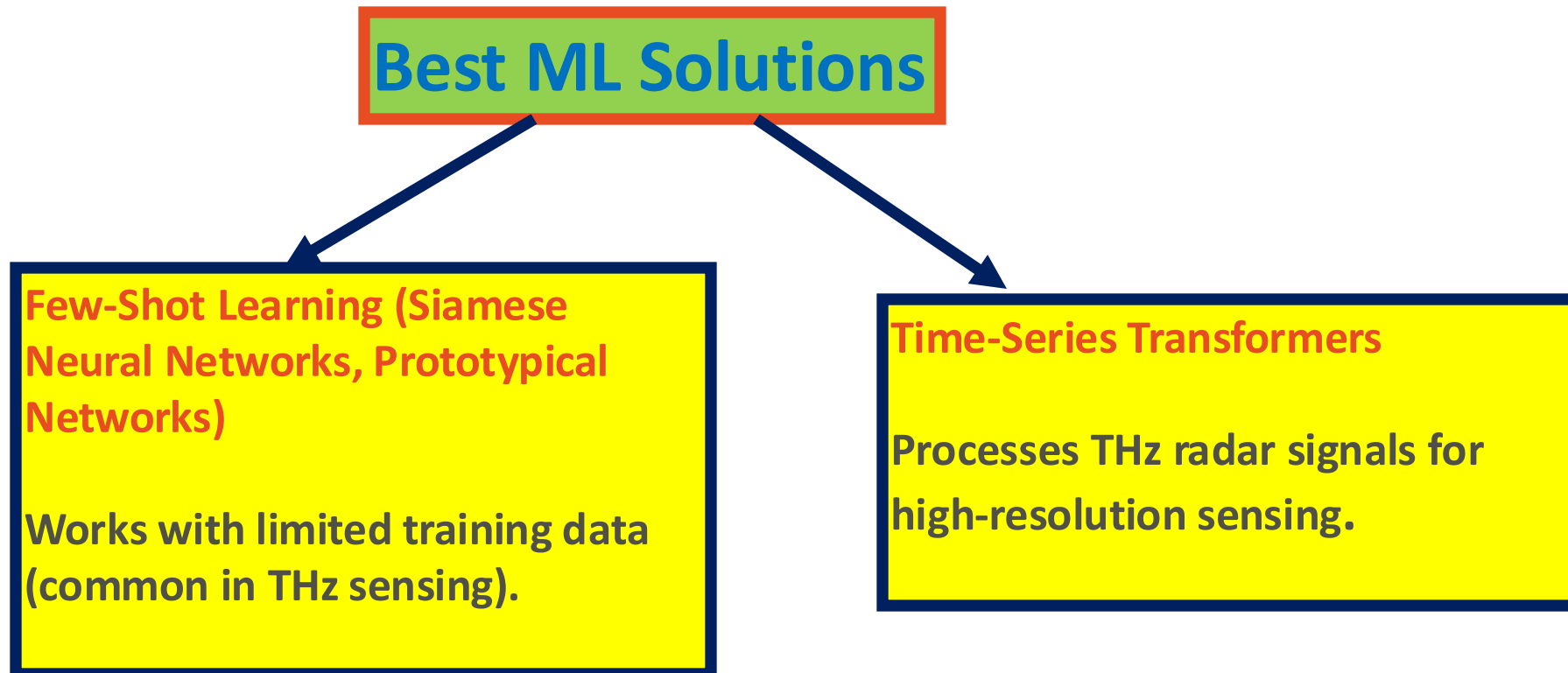


CASE 2: AI-Native THz BAND COMMUNICATION

5. THz Localization & Sensing

Problem:

THz enables mm-level localization, but requires high precision.



Future Research Directions for THz Band

FL for Distributed Systems

Enabling collaborative AI training across multiple THz BSs

Quantum Machine Learning (QML) for THz Signal Processing

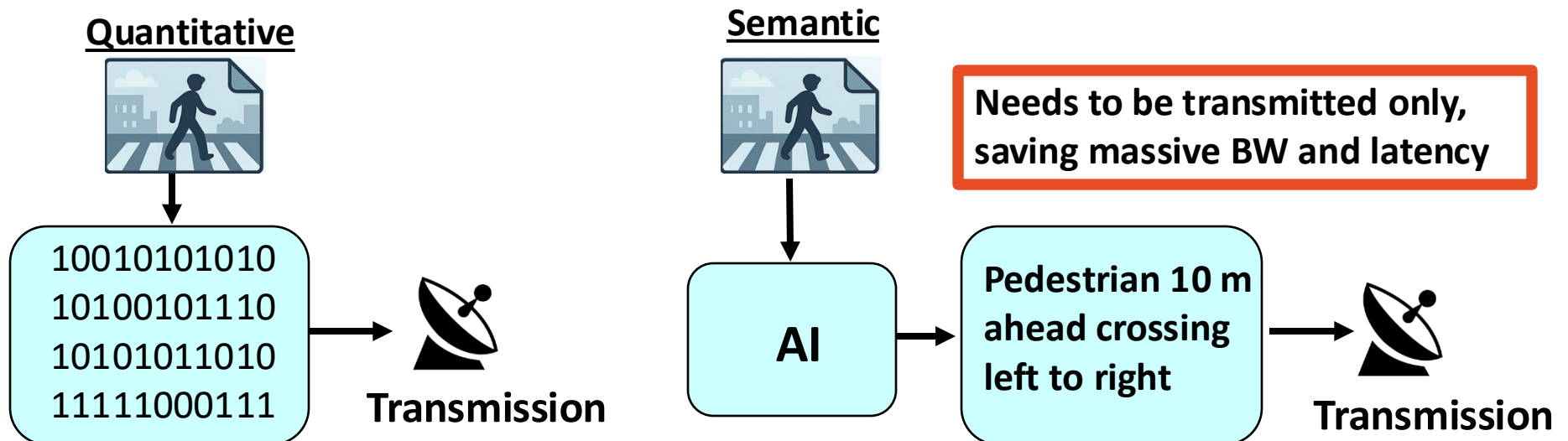
Leveraging quantum NNs for ultra-fast THz waveform optimization

Neuromorphic Computing for Low-Latency THz Beam Tracking

Spiking neural networks (SNNs) for energy-efficient real-time adaptation

Case 3: Quantitative vs Semantic (Qualitative) Communication

- **Quantitative Communication:** what is received = what is sent
- **Semantic (Qualitative) Com:** what is received = what is meant to send
 - Packets with small importance value can be dropped
 - Importance value can be determined by entropy



Semantic Communication

W. Weaver,

"Recent Contributions to the Mathematical Theory of Communication",

ETC: A Review of General Semantics, pp. 261-281, 1953.

Recent Contributions to
The Mathematical Theory of Communication

Warren Weaver

September, 1949



Claude Shannon



Warren Weaver

- **Technical Problem:**
How accurately can the symbols of communication be transmitted? (Shannon's Mathematical Theory)
- **Semantic Problem:**
How precisely do the transmitted symbols convey the desired meaning?
- **Effectiveness Problem:**
How effectively does the received meaning affect conduct in the desired way?

Case 3: How AI/ML Support Semantic Com in NG Wireless Systems

Semantic Information Extraction & Understanding

- Natural Language Processing (NLP) helps decode human intent in text/speech
- Computer Vision (CV) extracts high-level features (e.g., objects, actions) from images/videos
- Knowledge Graphs & Ontologies structure semantic relationships for efficient reasoning.

Context-Aware Transmission & Compression

- AI models (e.g., transformers, deep neural networks) identify and transmit only task-relevant information
- Semantic compression reduces data overhead by discarding irrelevant details (e.g., background noise in speech).

Case 3: AI-Native Semantic Communication in NG Wireless Systems

Adaptive Resource Allocation

- RL optimizes BW, power, latency based on semantic importance.
- Goal-oriented com prioritizes critical data (e.g., emergency alerts).

Semantic Channel Coding & Error Correction

- Neural decoders recover meaning even with errors.
- Generative AI (e.g., diffusion models) reconstructs lost semantic content.

Edge AI for Distributed Semantic Processing

- FL enables collaborative semantic understanding without raw data sharing.
- TinyML allows lightweight models on IoT/edge devices.

Key Challenges in AI-Native Semantic Communication

Interpretability & Trust

- Black-box AI decisions may lead to incorrect semantic interpretations.
- Need for explainable AI (XAI) in mission-critical applications.

Standardization & Semantic Frameworks

- No universal semantic encoding/decoding standards exist yet.
- Interoperability between different AI models is challenging.

Computational Overhead

Real-time semantic process. requires efficient AI models (e.g., quantization, pruning).

Privacy & Security Risks

- Semantic extraction may expose sensitive user context.
- Adversarial attacks could manipulate meaning (e.g., deepfake semantics).

Dynamic & Uncertain Environments

AI models must adapt to varying user intents and channel conditions.

Case 4: Zero-Touch Network Management (ZTNM)

Definition:

Self-configuring, self-optimizing, self-healing networks with minimal or no human intervention.

Why?

Human-managed networks cannot scale to billions of devices, ultra-low latency (zero latency), extreme reliability demands and managing zettabytes of data and critical energy consumption.

How?

Self-Healing Networks:

AI predicts and fixes failures before they occur (e.g., fiber cuts, congestion).

Intent-Based Networking (IBN):

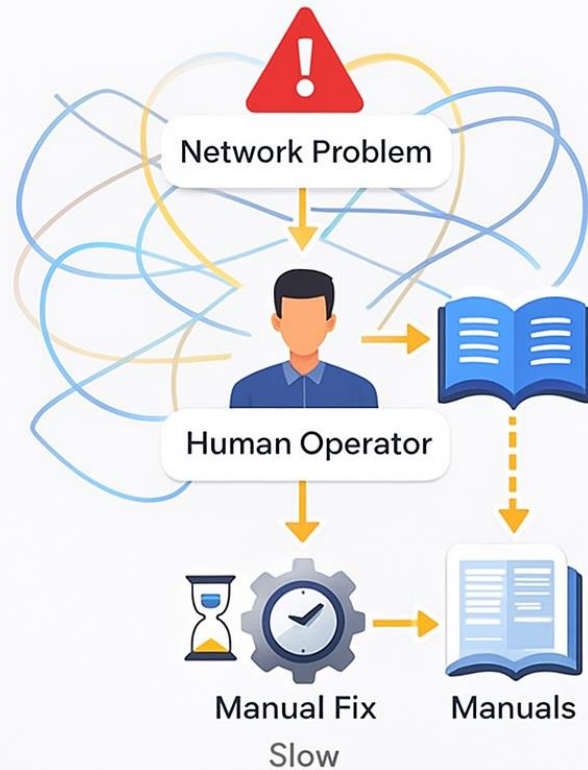
Operators define policies (e.g., "Ensure 1ms latency for AR/VR"), and AI executes them autonomously.

Automated SLA Enforcement:

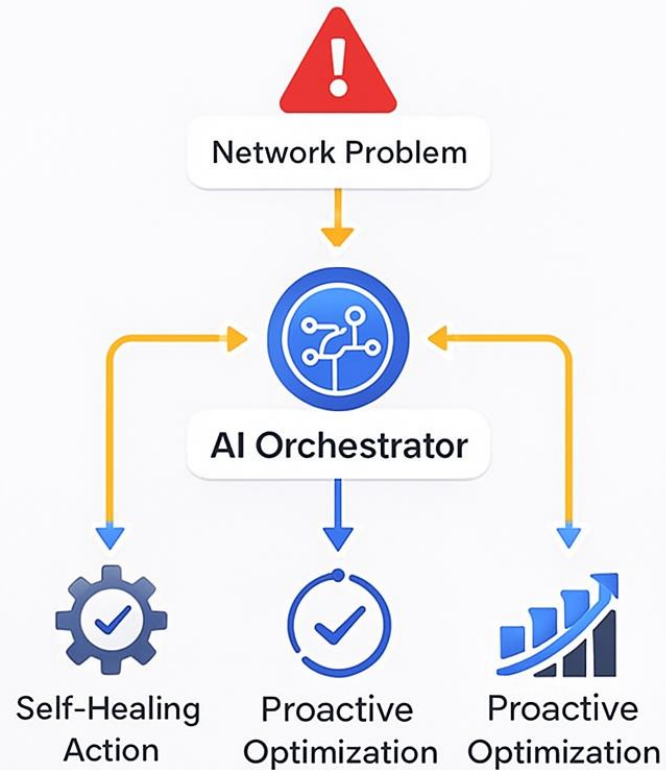
AI ensures QoS dynamically for different services (e.g., healthcare vs. gaming).

AI-Native ZTNM Solutions Deployed (Before and After AI Automation)

Before: Human-in-the-loop



After: ZTNM



AI-Native ZTNM Solutions Deployed (Before and After AI Automation)

	BEFORE	AFTER
OPEX Reduction (40%)	High labor costs for manual monitoring/troubleshooting	AI automates 80% of repetitive tasks
Fault Resolution Speed (30% Faster)	4-hour mean-time-to-repair (MTTR)	2.8-hour MTTR with AI-driven diagnostics
Energy Savings (20%)	BSs run at 100% power 24/7.	AI optimizes power usage (e.g., 50% power during off-peak)
Network Uptime (99.999%)	99.9% uptime (8.8 hours downtime/year)	5.3 minutes downtime/year

- Revenue Generation for Already Existing & New Services
- Reduce errors through minimization of human intervention
- Guaranteed adherence to dynamic and robust SLAs
- Increased Reliability and Efficiency

Key AI Technologies for ZTNM

- Deep Reinforcement Learning (DRL): For dynamic, real-time optimization.
- Federated Learning (FL): For privacy-preserving, collaborative AI.
- Digital Twins & Generative AI: To simulate and test networks using synthetic data.
- Explainable AI (XAI): To build trust and ensure decisions are transparent.

Grand Challenges on the Path to Autonomy

- **Data Problem:** How do we train AI without real 6G data? (Solution: Digital Twins, Federated Learning)
- **Trust Problem:** How do we trust a "black-box" AI? (Solution: Explainable AI)
- **Speed Problem:** How can AI make decisions in under 1ms? (Solution: Lightweight Edge AI)
- **Unity Problem:** How can AI manage a mix of satellites, drones, and ground system
(Solution: Cross-Domain AI Orchestration)

The Cognitive Wireless Era Vision: AI Based Unified Ecosystem

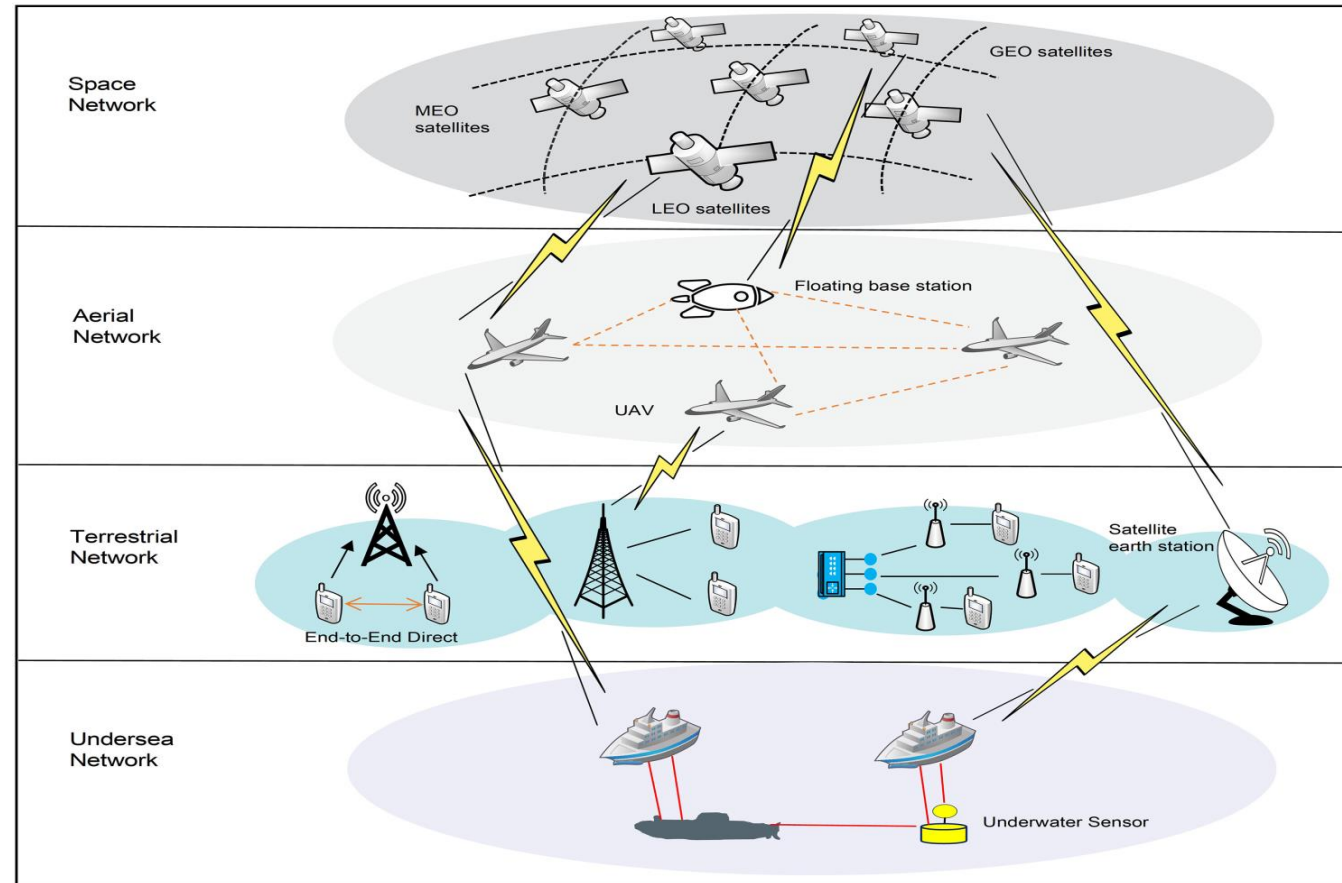
(Seamlessly Connecting the Deep Seas to Deep Space)

A single packet journeying from an UW sensor, to a drone, to a satellite, to a terrestrial BS, with an AI orchestrator managing the entire, complex handover in real-time. →
“The Cognitive Wireless Era”

- Each domain has different protocols (5G NR vs. acoustic modems vs. satellite DVB-S2X) & latency profiles (ms in 5G vs. sec in satellite vs. variable in UW,
- Manually managing handovers is impossible.
- Need a universal translator.

AI-based protocol translation (e.g., converting 5G packets to acoustic signals)

Challenge: No universal standards for AI-driven multi-domain networking.



**TOWER
OF
BABEL**

The Brain: The AI Orchestrator

- Centralized Intelligence:

Global view for optimal routing and slicing, in particular, manages cross-domain resource allocation, mobility & slicing.

- Distributed Agents:

Deployed at edge nodes for real-time decision-making.

- Federated Learning:

Enables collaborative AI training across domains without raw data sharing (without compromising data privacy) → synchronization is hard !!

Automated Network Management

AI automates network operations, reducing human intervention and improving efficiency.

- * Self-Healing and Self-Optimization

- Anomaly Detection:** AI detects and mitigates NW failures (e.g., satellite link drops, UW node malfunctions).

- Dynamic Reconfiguration:** AI adjusts beamforming, routing, and power allocation in real time.

- * Predictive Maintenance

- AI predicts HW degradation in UW sensors or satellite components using historical data**

- * Zero-Touch Network Management

- AI automates provisioning, scaling, and security policies across terrestrial, UW and space segments.**

AI-Driven Resource Allocation

Efficient resource allocation is critical in multi-domain networks due to varying latency, bandwidth, and reliability constraints.

Dynamic Spectrum Sharing

- AI optimizes spectrum usage across 5G, satellite, and UW acoustic bands.
- Reinforcement Learning (RL) adapts to interference and congestion.

Cross-Domain Load Balancing

- AI routes traffic through optimal paths (e.g., terrestrial vs. satellite backhaul)

Energy-Efficient Operation

- AI minimizes power consumption in underwater nodes and satellites.
- Sleep scheduling for IoT devices based on traffic prediction.

AI-Enabled Network Slicing

Network slicing creates virtualized, isolated sub-networks tailored for different services

Problem:

- Mission-critical communications (defense ops)
- Industrial automation (smart factories)
- Remote environmental monitoring (UW, space)
- Slices must dynamically adapt when switching domains (e.g., AUV moving from UW to satellite link)

AI Solutions & Challenges:

- DRL for slice lifecycle management
 - **Example:** AI reallocates resources when a submarine loses acoustic link and switches to buoy-satellite relay.
- **Challenge:** AI must guarantee isolation between slices (e.g., military vs. civilian traffic).

Security & Privacy in Multi-Domain Networks

Problem:

- Sensitive data (defense, offshore oil rigs, satellite imaging).
- Vulnerabilities increase when integrating:
 - UW sensors (easily tampered with)
 - Satellite links (susceptible to jamming)
 - Terrestrial edge nodes (physical attacks)

AI Solutions & Challenges:

- FL for secure model training
 - Keeps raw data localized (e.g., UW sonar data stays on AUVs).
- AI-driven intrusion detection
 - Detects anomalies in satellite-terrestrial handovers.
- **Challenge:** Adversarial AI attacks can fool detection systems.

Energy Efficiency & AI at the Edge

Problem:

- UW nodes (battery-powered, hard to recharge)
- Satellite payloads (limited solar power)
- Remote terrestrial edge servers (diesel generators in oil fields)

AI Solutions & Challenges:

- TinyML for low-power AI inference
 - **Example:** Lightweight AI models on UW sensors for anomaly detection.
- **Challenge:** Heavy AI workloads drain batteries faster.

Predictive & Seamless Handover Management

From Reactive to Proactive Mobility Management

- **Problem:** Blind handovers between high-speed LEO satellites and terrestrial cells cause packet loss and service disruption
- **AI Solution:** DRL for Handovers

Conclusion: The Cognitive Era is Here

- **Standing at the threshold of the cognitive wireless era.**
- **Shift from connecting things to connecting intelligence.**
- **From reactive networks to predictive, cognitive systems.**
- **AI-Native is the engine.**
- **Transformation of society, industry, and human potential.**
- **Question is no longer *if* this will happen, but how quickly we can responsibly build it.**

